Structural colours and applications to anodised aluminium surfaces

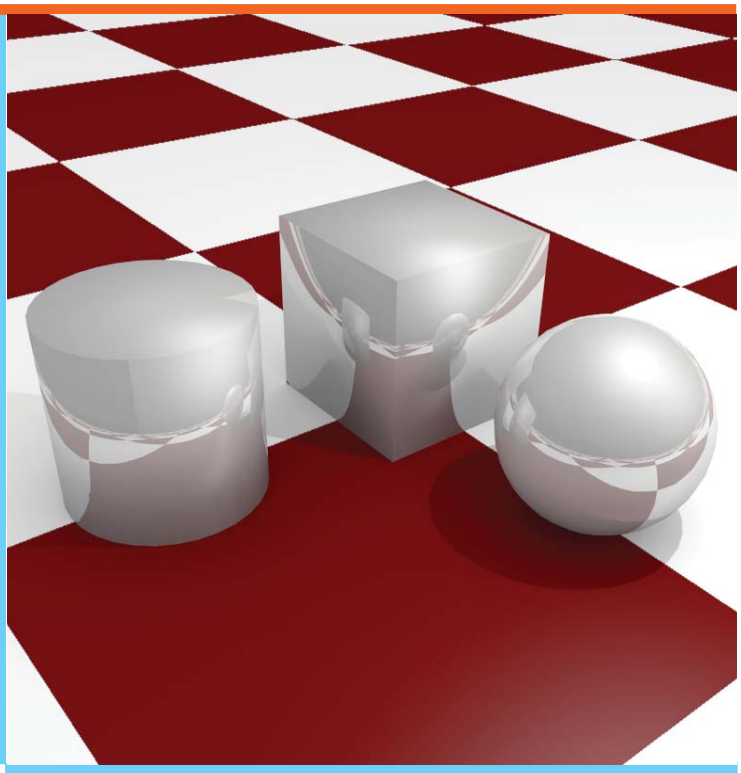
Johansen, Villads Egede; Sigmund, Ole; Aage, Niels; Breinbjerg, Olav

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Structural colours and applications to anodised aluminium surfaces

PhD Thesis



Villads Egede Johansen
DCAMM Special Report No. S173
November 2014

DTU Mechanical Engineering
Department of Mechanical Engineering

Structural colours and applications to anodised aluminium surfaces

by

Villads Egede Johansen

DTU Mechanical Engineering
Department of Mechanical Engineering

Title of the thesis:
Structural colours and applications to anodised aluminium surfaces

Ph.D. student:
Villads Egede Johansen
E-mail: vejo@mek.dtu.dk
E-mail: villads@egede.com

Supervisors:
Ole Sigmund
E-mail: sigmund@mek.dtu.dk

Niels Aage
E-mail: naage@mek.dtu.dk

Olav Breinbjerg
E-mail: ob@elektro.dtu.dk

Address:
Department of Mechanical Engineering, Solid Mechanics
Technical University of Denmark
Nils Koppels Allé, Building 404
2800 Kgs. Lyngby, Denmark

DCAMM Special Report no. s173
ISBN 978-87-7574-395-7

Whatever colors you have in your mind
I'll show them to you and you'll see them shine

Bob Dylan, Lay Lady Lay

Resumé (in Danish)

Følgende afhandling udforsker mulige metoder til at skabe aluminium med et mælkehvidt, metallisk udseende med dekorative formål for øje. Da farven hvid ikke kan opnås gennem normal absorptionsbaseret indfarvning af aluminium, er optiske mekanismer baseret på lysspredning fra nanostrukturer undersøgt for at løse problemet.

Til at starte med undersøges eksisterende arbejde inden for farver og visuel fremtoning. Dette inkluderer studier i hvordan farver opfattes og hvilke karakteristika man forbinder med en overflades udseende. Efterfølgende undersøges nanostrukturer af overfladeprofiler ved hjælp af optimering og topologioptimering, med det formål at klarlægge i hvilken grad farver kan frembringes ved manipulation af disse. Dette følges af en undersøgelse af hvilken betydning stokastiske variationer på nanoniveau har i forhold til at tilpasse overfladereflektioner og frembringe smukke farvefremtoninger. Bagefter gennemgås optiske modeller af lysspredende, ikke-deterministiske geometrier der kan bruges til at beskrive anodiseret aluminium.

Som en konsekvens af undersøgelserne præsenteres adskillige forslag til at opnå en hvid farvefremtoning i aluminium. Disse er beskrevet sammen med en kort forklaring om den aktuelle status for det eksperimentelle arbejde, der er udført for at teste forslagene. Resulater fra den mest succesfulde metode med hvilken det er lykkedes at opnå en hvid fremtoning er præsenteret til sidst.

Abstract

This thesis investigates possible ways of creating aluminium with a milky white, metallic appearance for decorative purposes. Since white cannot be obtained through traditional absorption based dyeing of aluminium, optical mechanisms based on scattering by nanostructures are studied in order to solve the problem.

The problem is investigated by first reviewing existing work within colouration and visual appearance. This includes a study on how colours are perceived by humans and an investigation of the characteristics with which a surface appearance is properly described.

Subsequently, nanostructures and surface profiles are investigated using optimisation and topology optimisation in order to understand the limitations and design freedom of colour engineering. This is then followed by a study of the effect of disorder on a nanoscale level in order to tailor surface reflections for a smooth, pleasing appearance. Afterwards, optical models for scattering of non-deterministic geometries suitable for anodised aluminium are considered.

The outcome of the investigations are several different proposals for obtaining a white appearance for aluminium. These are described in the thesis alongside a brief note on the experimental work performed to try out the proposals. Results from the most successful experiment in which a satisfactory white appearance is obtained is then presented.

Preface

This thesis is submitted in partial fulfilment of the requirements for obtaining the degree of Ph.D. in mechanical engineering at Technical University of Denmark (DTU). The Ph.D. project was funded by the Danish National Technology Foundation (HTF, Højteknologifonden) through the project Optically Designed Anodised Aluminium Surfaces (ODAAS). It was carried out at the Department of Mechanical Engineering, Solid Mechanics (FAM) at DTU in the period 1st December, 2011 – 30th November, 2014. Supervisors on the project were Professor Dr. techn. Ole Sigmund, Associate Professor Niels Aage and Professor Olav Breinbjerg.

I would like to thank my supervisors Ole Sigmund for his excellent guidance and intuition during the project, Niels Aage for his helpful attitude and collaboration, and Olav Breinbjerg for a thorough introduction to electromagnetism. Special thanks are given to Professor Peter Vukusic for kindly allowing my research stay at University of Exeter and for fruitful discussions and feedback on my work. Also, I owe thanks to Associate Professor Akira Saito for accepting my visit at Osaka University which was a great source of inspiration for me. Furthermore, Professor Jung H. Shin and Professor Bill Barnes are acknowledged for their proposals leading to the extra investigations in Section 6.3 and the optical brightening approach in Section 8.1 respectively.

I am grateful for having collaborated with my cheerful and knowledgeable colleagues in the ODAAS project. This includes former Ph.D. student Martin Aggerbeck who has been a big help and a good discussion partner. Furthermore, it has been a pleasure to share an interest in optics with Ph.D. Jacob Andkjær both while he was employed at FAM and afterwards. I am also thankful for the satisfactory collaboration with NIL Technology ApS which ended up with the sample produced in P4.

It has been a true joy being around my colleagues at FAM and in the TopOpt group. Especially my next-desk neighbour Erik Andreassen has been a gratifying professional and personal conversation partner. At University of Exeter, my stay was made enjoyable by the basement team: Luke, Matt, Tim, Dan and Caz (they may have moved your desk, but your heart belongs to the basement). Also, the G31's and Tom gave me a pleasant time at Exeter.

I am grateful for having friends that support a life full of other healthy interest than engineering. In particular I find great support in Johan. I have furthermore enjoyed the years I have spend in Kollektivet Hegnet. Also, I enjoy ever well-spend time playing football with my friends from FC Salsa Akademiet.

I am thankful for the joy and support my partner Marta is able to give. Lastly, I am grateful for having a family who supports me and listens to me no matter what I am doing.

Villads Egede Johansen,
30th November, 2014

Publications

The following international journal publications are a part of the thesis:

(P1) V. E. Johansen, J. Andkjær and O. Sigmund. Design of structurally colored surfaces based on scalar diffraction theory, *Journal of the Optical Society of America B*, **31**(2), pp. 207–17, February 2014.

(P2) J. Andkjær, V. E. Johansen, K. S. Friis and O. Sigmund. Inverse design of nanostructured surfaces for color effects. *Journal of the Optical Society of America B*, **31**(1), pp. 164–74, January 2014.

(P3) V. E. Johansen. Optical role of randomness for structures surfaces. *Applied Optics*, **53**(11), pp. 2405–15, April 2014.

(P4) V. E. Johansen, L. H. Thamdrup, K. Smistrup, T. Nielsen, O. Sigmund and P. Vukusic. Designing visual appearance using a structured surface. *Optica (accepted).*

(P5) V. E. Johansen. Preparing the Generalized Harvey-Shack rough surface scattering method for use with the Discrete Ordinates Method *Journal of the Optical Society of America A*, **32**(2), pp. 186–94, February 2015.

(P6) V. C. Gudla, S. Canulescu, V. E. Johansen, J. Schou and R. Ambat. Reflectance spectroscopy of polyurethane-TiO_2 composites on aluminium. *Journal of Materials Science (to be submitted).*

The following journal publication and conference proceeding are published during the Ph.D. project but does not constitute a part of the thesis:

(P7) M. Aggerbeck, S. Canulescu, K. Dirscherl, V. E. Johansen, S. Engberg, J. Schou and R. Ambat. Appearance of anodised aluminium: Effect of alloy composition and prior surface finish. *Surface and Coatings Technology*, **254**, pp. 28–41, September 2014.

(P8) VC. Gudla, S. Canulescu, VE. Johansen, J. Schou and R. Ambat. Reflectance spectroscopy from TiO2 particles embedded in polyurethane. *Proceedings of EUROCORR 2013*, September 2013.

Contents

1 Introduction

There is a crack in everything,
that's how the light gets in

Leonard Cohen, Anthem

This thesis is a part of the project Optically Designed Anodised Aluminium Surfaces (ODAAS) funded by the Danish National Technology Foundation (HTF, Højteknologifonden) and the high-end electronics producer Bang & Olufsen (B&O). The primary aim is to create aluminium with "*an anodised white (milky) appearance*"[1] to extend "*B&O's design possibilities beyond the current limitations*"[1]. Currently, aluminium can be fabricated in a wide variety of colours by adding dye to anodised aluminium before sealing it, see Figure 1.1(a). White is not one of these colours since it cannot be obtained through a dyeing process.

When introduced to the project, people often ask why B&O is not satisfied with having the surface painted? The answer is that aluminium is a characteristic feature of B&O designs, and when a metal is spray painted, it does not look like a metal any more. Instead it looks like plastic or ceramic (just like the appearance of most fridges). This difference is captured in Figure 1.1(b).

The market potential for a white appearance spans beyond B&O's products since white is a trendy colour and other high-end products might benefit from it as well. Besides other decorative purposes, the material and appearance combination will work well for medical instrumentation where a white colour and an easy-to-clean surface serve clinical purposes.

1.1 ODAAS

The ODAAS project is structured such that a majority of the participants work with processing and analysis of aluminium. These are researchers from the Section of Materials and Surface Engineering at the Technical University of Denmark (DTU) as well as employees at B&O and the Tribology Centre at the Danish Technological Institute (DTI). Besides this large production and fabrication track, the project is supported by participants performing optical modelling and characterisation in order to predict and discard methods for obtaining the desired aluminium appearance. The optical characterisation is carried

[1] Quote from the original project description.

(a) (b)

Figure 1.1: (a) A B&O telephone in different anodised aluminium colours – except the white one which has been spray painted. (b) Close-up of the red and white B&O telephone shells: notice the difference in reflection from the lamp in the red, metallic surface compared to the white, ceramic looking surface.

out by participants from the Optical Microsensors and Micromaterials group from DTU Photonics (Fotonik) at the Risø DTU campus, and the optical modelling is carried out at the Section of Solid Mechanics (FAM) at DTU. The Ph.D. project linked with this thesis is the contribution from FAM to the project.

1.2 Surface treatment of aluminium

Aluminium is the most abundant metal in the Earth's crust. It has a remarkable low density, good corrosion resistance and different aluminium alloys cover a wide range of mechanical properties. For these reasons, aluminium finds use for everything from thin household aluminium foil to cars, aerospace and construction. Since melting aluminium only requires 5% of the energy needed for extracting aluminium from the Earth's crust (Aggerbeck, 2014), recycled aluminium constitutes an increasing part of aluminium products. However, the recycling process adds impurities to the aluminium which in turn influences its optical, mechanical and corrosion properties.

The choice of surface treatment(s) of aluminium has a predominant influence on its visual appearance. Methods like bead blasting and etching gives a rough surface with matt appearance whereas diamond turning or wheel polishing is used for highly glossy/reflective surfaces. Processes like degreasing and desmutting remove contaminants from the surface. The mentioned treatments need to be taken into account when considering surface appearance, yet the widest range of possibilities to alter aluminium's visual appearance is by *anodisation*. This is briefly described in the following Section.

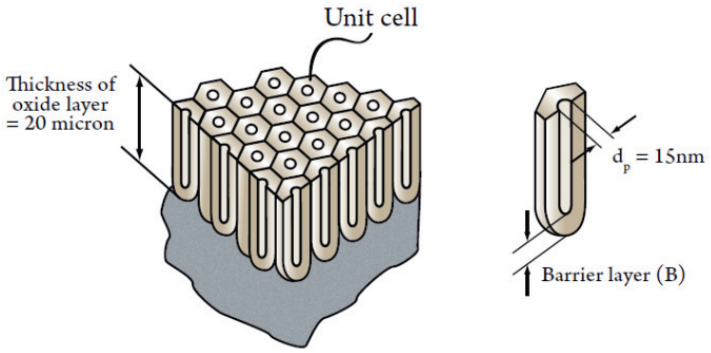


Figure 1.2: When aluminium is anodised, pores of alumina (Al_2O_3) grow on the surface.

Anodisation

Anodisation is a process used to increase the natural oxide layer thickness on the surface of metals. It consists of applying a voltage difference between the metal (used as an anode, hence the name of the process) and another electrode (the cathode) dipped into a bath of, for example, sulphuric acid. Having a micrometer thick anodised layer increases corrosion and wear resistance and alters the visual appearance of the surface. Most aluminium parts produced by B&O are anodised.

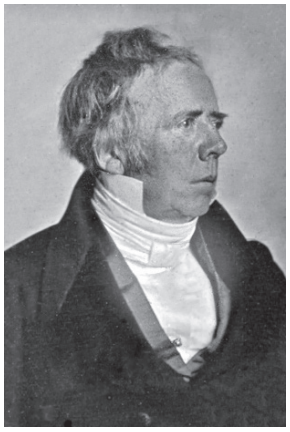
Aluminium oxide is transparent with a refractive index around 1.6-1.7, and has the peculiar property that regular pores can be formed during anodisation, see Figure 1.2. After anodisation, the oxide is usually sealed by dipping it in hot water, which fills and collapses the pores. It is possible to fill the pores with dye before this step, which is the method used to create the colours in Figure 1.1(a). Pigments normally used for white paints are a magnitude larger than the oxide pore diameters.

Much of the experimental work in ODAAS focuses on controlling the anodisation processes in order to investigate the possibilities anodisation gives with respect to aluminium appearance.

An in-depth description of treatment of aluminium with focus on appearance is found in Aggerbeck et al. (2014a) and Aggerbeck (2014).

1.3 Hans Christian Ørsted

It shall not go unnoticed that one man – from a historical perspective – has laid an important part of the foundation for this work, namely Hans Christian Ørsted (see Figure 1.3). Not only was DTU founded on the initiative of Ørsted, he was at the same time the first scientist to isolate pure aluminium. As if this cornerstone contribution of discovering the metal to which all effort in ODAAS is put into is not enough in itself, Ørsted also discovered a direct relationship between electricity and magnetism. This is a substantial contribution to the later development of the theory on electromagnetism through which the behaviour of light can be described. It is remarkable how work from nearly two centuries ago still joins in unity in novel ways to this day.



(image: Wikimedia)

Figure 1.3: Hans Christian Ørsted (1777–1851).

1.4 Reading guide

The thesis is structured such that it first gives an introduction to **Colours and visual appearance** as a motivation to the subject as well as to make the reader reflect upon the subject of colours and appearance from a scientific perspective. This is followed by a chapter on **Rendering and visualising appearance** which describes appearance in more technical terms. The chapter introduces knowledge from computer graphics research, where visual appearance has been studied for decades, and it ends up by giving two examples to motivate that surface appearance is not just described by a colour – but by its bidirectional reflectance distribution function (BRDF). Since light is described by waves, a chapter on **Wave properties of light** follows, which gives an introduction to central mathematical concepts used in the thesis to calculate the BRDF of a given surface.

These introductory chapters are then followed by several chapters on analytical work carried out during the project in order to grasp the possibilities within appearance design as well as to obtain understanding of visual appearance analysis. This is initiated by considering how **Optimised structural colour designs** are designed and the implications thereof. Subsequently, a chapter devoted to the influence on nanoscale structural **Order and disorder** investigates the significance of stochastically described structures to visual appearance. After this, a chapter on **Appearance of aluminium** presents methods to describe effective scattering properties of stochastically described structures with special focus on their application to aluminium.

The accumulated knowledge from the just mentioned research is then used to generate a range of ideas for **Creating white aluminium**. The experimental work on those ideas are evaluated as well, and it is shown how a white aluminium appearance is obtained. The work is summarised by a few **Concluding remarks**. After the the appendices, a collection of articles which constitutes the thesis is found.

2 Colours and visual appearance

Yellow is the colour of my true love's hair
In the mornin' when we rise

Donovan, Colours

The sensing and signal processing capabilities of humans are fantastic beyond imagination! Take a minute to be amazed by the fine mechanical sensing of our ears; the chemical detection in our nose and tongue; the surface profile, pressure and temperature measurements when touching; and the terahertz electromagnetic sensing of our eyes.

Notwithstanding the great pleasure lying in using our senses, this chapter is devoted to a short description of one of them, namely vision.

2.1 Light

What the eyes detect is called light. The behaviour of light can in many cases be treated as *rays*, but when it interacts with structures comparable to the wavelength of light, it has to be treated as electromagnetic *waves* that fulfils Maxwell's equations. This enables prediction of more complex behaviour than ray analysis. For small quantities of light, Maxwell's equations break down, and the behaviour follows the rules of quantum electrodynamics (QED) where light is treated as photons (Feynman, 1988). These three different views on light are – in the order mentioned – generalisations of each other, and can be sought to explain increasingly complicated features of the behaviour of light, see Figure 2.1. Wave and ray theories adequately describes light interaction with structures presented this thesis, and this will be the basis of analysis in the subsequent chapters.

2.2 Visual appearance – the *perception* of colours

Often objects are classified and differentiated by verbalising their appearance based on their colour. This, for example, makes it possible to estimate if a banana is ripe just by evaluating its colour. Such a classification is useful in most practical cases, but at the same time insufficient to fully describe an object or a material. The approach is inaccurate for example in case of transparent (water, oil) or iridescent (soap bubbles, oil puddles,

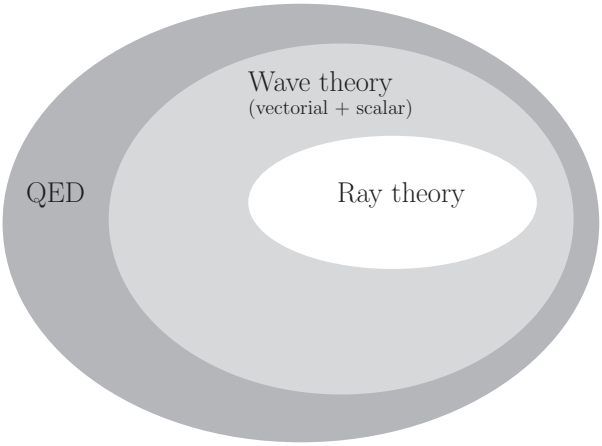


Figure 2.1: Different theories to explain the behaviour of light. QED is the most general theory. The wave theory of light considers the average behaviour of large ensembles of photons described by QED. Ray theories are valid when light waves only interacts with structures much larger than the wavelength of light.

(image: Martin Aggerbeck)

Figure 2.2: Different kinds of white: even though all objects can be described as white, they differ in gloss, brightness and hue. None of these give the feeling of white metal.

CD's) materials. This is because materials and surfaces are strongly dependent on the environment in which they are observed: some might depict mirrors as grey, but in reality mirrors take on the colour of the objects they reflect; white walls look white even under dim light where they should appear grey; and if all light is turned off in a room, all objects appear black.

Other surface characteristics are equally important for the perception of a surface. Is the surface glossy or matte? Is it woven? Has it been polished? All these features add to the *visual perception* of an object, thus explaining why painting aluminium is not an adequate solution for B&O: in that way it is not *perceived* as white aluminium. In Figure 2.2 different white objects are compared to exemplify the importance of surface characteristics.

The purpose of this chapter is to exemplify how colours in themselves inadequately describe an object's visual properties. When designing, it is in many cases relevant to consider all factors contributing to visual appearance rather than only the perceived colour.

Figure 2.3: Cut out the two holes indicated by dashes. The optical illusion of different shades of orange in Figure 2.4 will vanish when the colours are observed freed from their environment.

Misperceptions

A simple example of colours failing to adequately describe an object is seen in Figure 2.4. The two orange circles are of exactly the same colour, but perceived differently due to the surrounding environments. By following the instructions in Figure 2.3, the objects can be observed in similar environments, thus revealing their identical colour to the eye.

The reason for the optical illusion, is that interpretation of colours is the combined work of the receptors in the eye and cognitive processing of that information. This is referred to as chromatic adaptation or colour constancy. The effect makes absolute estimates of visual appearance of a quantified/measured colour difficult. From a design perspective, it also gives the possibility to trick the eye in order to obtain a desired appearance.

It should be kept in mind that all attempts to visualise colour reflections throughout the thesis are indicative approaches. The actual visual appearance of a fabricated theoretical structure can vary significantly. If an appearance should be represented appropriately, it has to be rendered in a proper environment using computer graphics.

2.3 Colours

Now that some pitfalls of attributing colours to objects have been presented, it is appropriate to introduce knowledge on how colours can be perceived. This section explains how the electromagnetic light spectrum is converted into a colour and which consequences it has. This will form the basis for investigating visual appearance later on in the thesis.

The eye

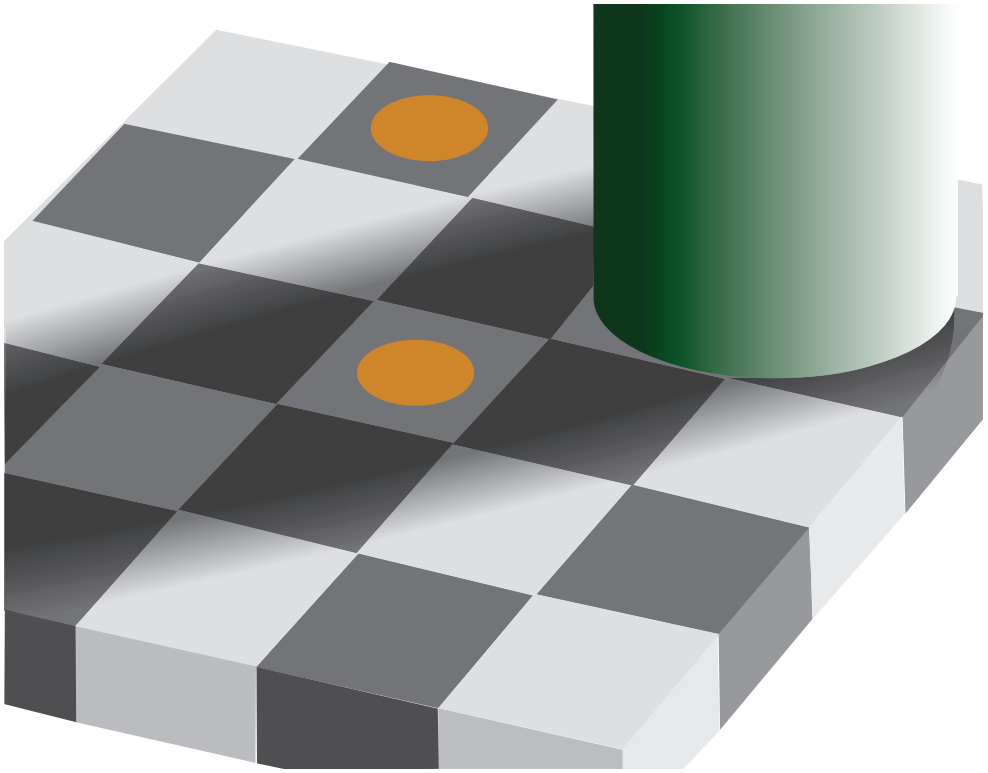
To completely understand the conversion of light energy to colour, knowledge on optical systems, detectors, neural processing and cognition is required (Billmeyer and Saltzman, 2000). For the purposes in this thesis, knowledge on the conversion of light spectra into colours is sufficient to construct an adequate framework in which colours and appearance can be treated.

Light energy is converted to colours by the eye. The eye consists of two classes of receptors, *rods* and *cones*, where the rods detect small light intensities such as starlight – but no colours. At higher intensities the rods stop sending signals to the brain, and the cones take over. There are three different cones with peak sensitivities in different parts of the visible spectrum. By processing these stimuli in the brain, light intensities can be converted into colours. Since an understanding of all the mechanisms in the brain and the eye is immensely complex, experiments like the one described in the following section have been used to map the effective behaviour of this system.

Colour transformations and the CIE experiment

In 1931 it was agreed upon by the Colorimetry Committee of Commission Internationale de l'Eclairage (CIE) to adopt three so-called colour-matching functions for a standard observer as the accepted way of converting light spectra into colours (Fairman et al., 1997). The functions are based on an experiment carried out in 1929 (Wright, 1928–29) based on 36 observers. The functions are still in widespread use today. Corrections and other functions have since been developed, but the differences between the functions are too small to have significant impact on the results in the thesis.

To convert a light intensity spectrum $I(\lambda)$ to a colour in the RGB (Red, Green, Blue) colour coordinate system, the following integrals are carried out:



(image: Wikimedia)

Figure 2.4: An optical illusion: the two orange discs (and the squares behind) are of exactly the same colour. They are perceived differently due to the environment they are appearing in. Observed free of the scene, the colour similarity is easily noticed, see Figure 2.3.

$$R = \int_{\infty}^{\infty} I(\lambda)\bar{r}(\lambda)\mathrm{d}\lambda, \tag{2.1a}$$

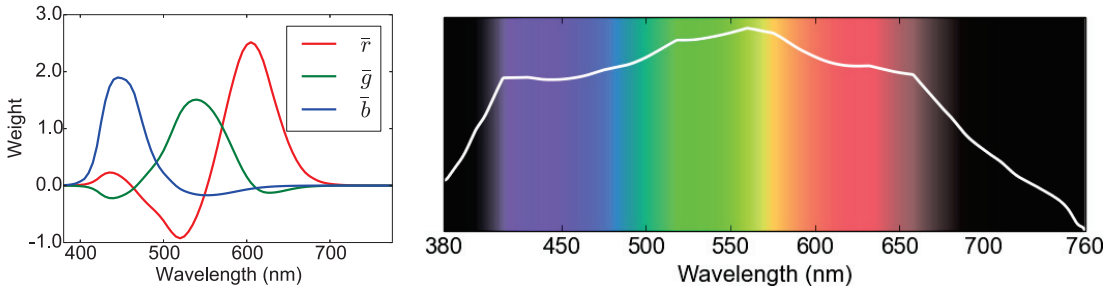
$$G = \int_{\infty}^{\infty} I(\lambda)\bar{g}(\lambda)\mathrm{d}\lambda, \tag{2.1b}$$

$$B = \int_{\infty}^{\infty} I(\lambda)\bar{b}(\lambda)\mathrm{d}\lambda, \tag{2.1c}$$

where $\bar{r}, \bar{g}, \bar{b}$ are the colour matching functions, see Figure 2.5(a). Tabulated values can be found online at `cie.co.at` or in Billmeyer and Saltzman (2000). Colours are adequately described in a three-dimensional space since the eye has three different cone receptors as described earlier. For compact notation of colours, RGB triplets will be represented in vector notation such that

$$\boldsymbol{C} = (R, G, B)^T. \tag{2.2}$$

Many other colour coordinate systems exist such as XYZ (old system, but easy to obtain from measurement); CIELab (close to perceptionally uniform, and intensity isolated



(a) (b)

Figure 2.5: (a) RGB colour matching functions. (b) Representation of the visible light spectrum generated using the colour matching functions followed by a contrast/brightness adjustment. The dark areas are not necessarily colourless, but a much larger relative intensity will be needed to generate a visible violet/red colour. A measure of colour lightness from a so-called CIElab conversion is given by the white curve. The order of the colours follow the order in a rainbow.

along one axis); and HSB (colours expressed by their hue, saturation, brightness). All systems have their own application, and, sometimes non-linear, conversions between different representations are possible. Research in colour spaces is still active (Li and Lee, 2014).

The visible light spectrum

Referring to (2.1), each wavelength in the visible light spectrum can be associated with a colour. By generating these colours, a rainbow plot is obtained as seen in Figure 2.5(b). This rainbow palette does not cover all colours since spectra often consist of a wide range of wavelengths. It is still a good guideline as to what intensities are dominating if a blue/green/red/... colour is observed. From the overlaid curve in Figure 2.5(b) it is seen how the eye is almost equally sensitive to all wavelengths in a wide range, albeit with a slight favour for green light around 550 nm.

Metamerism

An interesting consequence of the colour conversion is that the same colour may arise from different spectra. This is in line with what in colourimetry is referred to as metamerism, which is the matching of apparent colours from objects with different light spectra. It is this effect that makes it possible to display white on a computer screen just by combining red, green and blue light appropriately. It is for the same reason that we feel the quality of light is worse from a fluorescent lamp – that has many abrupt changes in its spectrum – than a traditional incandescent bulb or a halogen lamp with their more soft spectra: even though all three light sources appear similar, colour reflection from surfaces are, of course, best evaluated if the whole visible light spectrum is present.

Colour theory

In painting and visual arts, colour theory[1] is used as a guidance for colour mixing and analysis[2]. Evidence of trying to classify colours and deduct their properties goes back to the great Italian artists Leone Battista Alberti and Leonardo da Vinci (15th century). Through time, great minds such as Isaac Newton, Thomas Young, James Clerk Maxwell and Hermann von Helmholtz have also been involved[3]. None of these great personae were able to create a theory that can serve as more than guidelines for creation of visual arts.

As explained, this is because there is more to visual appearance than just quantification of colours. Still, these theories may serve as useful tools for designing and understanding colours. A few highlights of colour theory are presented in the following as an aid for understanding and further develop the concepts encountered throughout the thesis.

Primary colours Primary colours are a set of (normally three) colours which can be used to generate a wide palette of colours by combining them. The reason why three colours are needed is that the three different cones in the eye each need to be stimulated individually. A set of primaries will most likely not cover the whole range of visible colours, but a good part of them.

Subtractive colour mixing ("by pigment") When painting or printing, colours are created by adding pigments that absorb a part of the light spectrum. The primary colours here are cyan, magenta and yellow (CMY). From a white background (the whole visible spectrum), it is then possible to construct a wide colour palette by combining these three primary colours. The principle is illustrated in Figure 2.6.

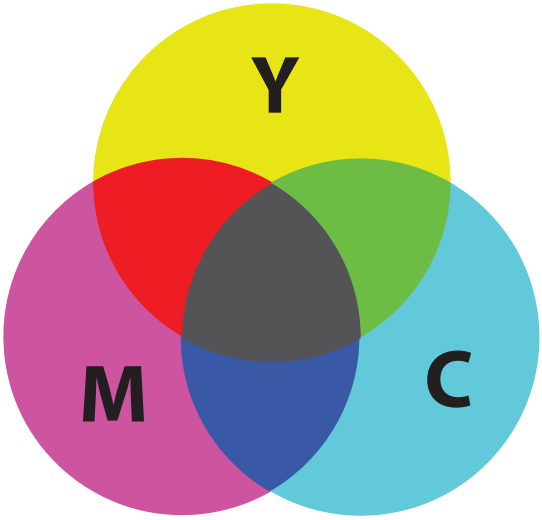
In the practical case of printers, a separate black colour ink is normally present in order to save ink and get as black a colour as possible. This is denoted "key". Thus, when printing, the CMYK colour model is used.

Additive colour mixing ("by light") Additive colour mixing can be seen as the opposite of subtractive colour mixing. Here, colour is created by adding light intensity (for example by using light sources). The primary colours here are red, green and blue (RGB), and black (no light intensity) is then needed as a background colour, see Figure 2.6. A well-known example of additive colour mixing is in TV and computer screens, where RGB pixels which can be varied in intensity are packed so closely, that they are observed as one point with the desired mixed colour.

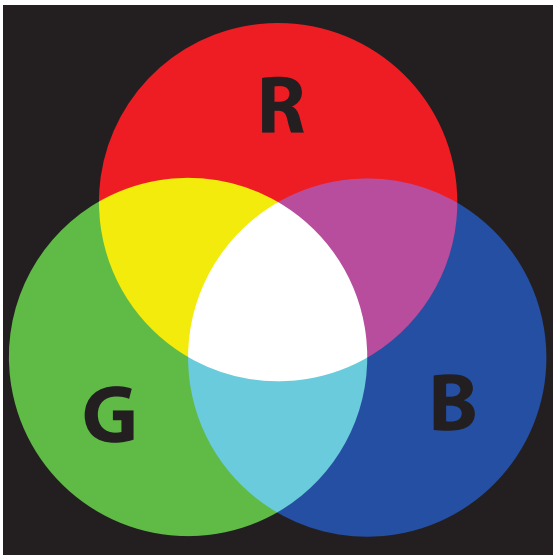
[1] For Danish readers, please note that "farveteori" is not in itself a technical term. We refer to the concepts mentioned as "farvesystemer" or "farvelære".

[2] Such guidances may be deliberate choice of complementary colours or warm and cool colours.

[3] Oliver Heaviside is conspicuous by his absence. Alas, the topic might have been too colourful for him.



(image: Wikimedia)
Subtractive colour mixing



(image: Wikimedia)
Additive colour mixing

Figure 2.6: Principles of colour mixing.

2.4 Visual appearance mechanisms

The mechanisms behind visual appearance of surfaces and optical phenomena are in general many and complex. To give the reader an idea of these, a few everyday examples of the marvellous world of colours are presented in Table 2.1 below. The book Bohren (2001) is recommended reading for educational, everyday experiments on optical phenomena. Many of the principles will occur later on in the thesis.

(image: Wikimedia)

Absorption The most common way of colour creation is by absorption – like in a paint or dye. Depending on particle composition, some materials happen to absorb light and therefore only reflect a part of the white light spectrum. A blue paint, for example, is absorbing the red and green part of the visible spectrum (due to subtractive colour mixing).

(continues on next page)

Table 2.1: Descriptions of different visual appearance mechanisms.


(image: Wikimedia)

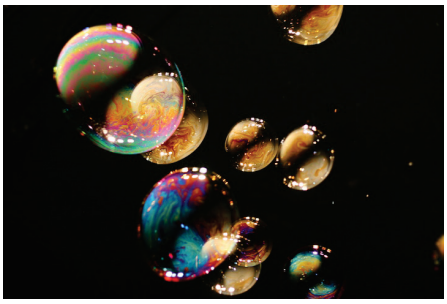
Thermal radiation All objects emit thermal radiation, which is electromagnetic radiation generated by the thermal motion of charged particles. This radiation might be within the visible light range, and that is why hot objects (like melted metal or a flame) appear red or yellowish. This is the working principle of incandescent light bulbs and the Sun. Light can also be generated by stimulated or spontaneous emission.

Small particle scattering Very small particles (smaller than the wavelength of light) scatters blue light more than red light since the wavelength of blue light is shorter than that of red light. Particles therefore appear optically larger for blue light. Even though the individual particle does not scatter much light, huge ensembles of such particles will give rise to blue light being scattered more than red. This is why the sky is blue (scattered sunlight) and red only at sunset (more light from the forward scattered direction). This is called Rayleigh scattering.

Particle scattering I For larger particles, scattering becomes more complex, and all wavelengths are strongly scattered. Imagine a setup where light shines through an ensemble of such scatters onto a canvas. This will enable visualisation of which directions different wavelengths are scattered. This is exactly what happens with rainbows, where the atmosphere acts as a canvas. For this reason, the Sun is always in your back when you are facing a rainbow (note the shadow of the traffic sign on the picture). This means that wavelength dependence is still present, but it is much more difficult to predict than for small particles.

(continues on next page)

Table 2.1: Descriptions of different visual appearance mechanisms.

	**Particle scattering II** Large ensembles of larger particles scatter light at all wavelength so much, that it in practice looses its directionality. This is how white paint is created and also why milk looks white (the fat particles act as scatters).
 (image: soapbubble.dk)	**Thin film interference** When light is reflected back an forth in thin transparent materials it experiences extraordinary reflection and transmission depending on wavelength. This means that such thin films act as colour filters. The principle is seen in soap bubbles and oil on water puddles.
 (image: Wikimedia)	**Diffraction** If a high degree of nanoscale order is present, light reflection will be strongly wavelength dependent, meaning that different wavelength in general will be scattered at different angles. This leaves an impression of a surface with colours sensitive to viewing angle. The effect is prominent for a compact disc (CD), where data is stored in a highly regular pattern.

(continues on next page)

Table 2.1: Descriptions of different visual appearance mechanisms.

(image: Wikimedia)

Retroreflection Retroreflectors are devices that reflect light strongly back towards their source. Retroreflectors are often used for traffic sign and visibility clothing. They are also found in nature, for example in the eyes of cats. The working principle of retroreflectors is to ensure a geometry which, for a wide range of incident angles, can reflect light back in the same direction it entered by multiple reflections. A typical retroreflector geometry is the so-called corner reflector. Apollo 11 brought a retroreflector to the moon for laser ranging experiments.

Table 2.1: Descriptions of different visual appearance mechanisms.

Structural colours

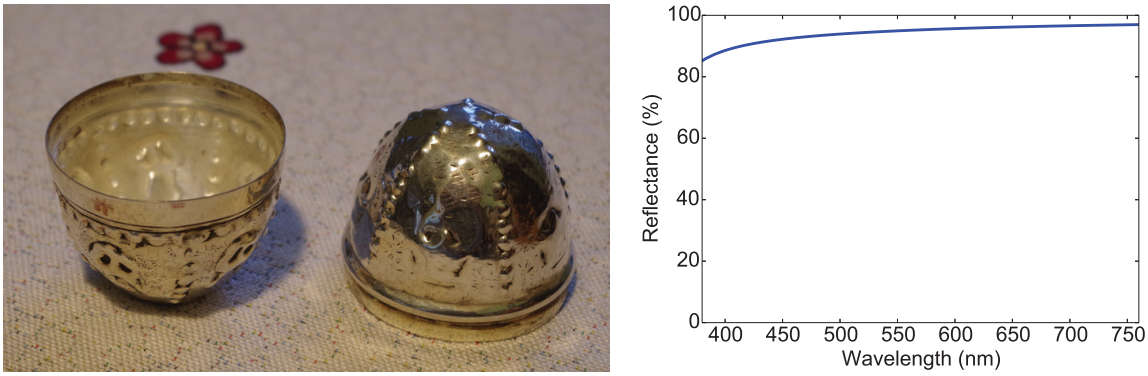
In this we are able from a colourless body to produce several coloured bodies, affording all the variety of Colours imaginable

Robert Hooke, Hooke (1665).

When the structures that light interact with are comparable to the wavelength of light – as is the case for many examples in Table 2.1 – light can interact with materials in completely different ways than otherwise observed. It is for example surprising how soap bubbles, which are made of a transparent material (soap water), all of a sudden exhibit a vivid colour appearance. Such extraordinary light interactions due to micro-/nanostructures that gives rise to colour appearances is called *structural colours*. Most discoveries within structural colour research come from detailed studies of animals, plants and bacteria. Structural colours have existed for at least 500 million years in nature (Parker, 2000). Whether mentioned explicitly or not, most of the work presented in this thesis falls under the category of structural colours.

2.5 White in particular

The colour white deserves special attention since the main purpose of the thesis is to find ways to modify the appearance of an aluminium surface such that it will be perceived as white aluminium.



(a) (b)

Figure 2.7: (a) Which is whiter? The polished outside or opaque, non-polished inside of this silver egg? And which look more metallic? Both reflect most of the visible light spectrum but in different ways. (b) Normal incidence reflection spectrum for silver (data from `refractiveindex.info`).

Whereas most colours can be produced by absorption, this is not the case for white, since it is a combination of all visible wavelengths of light. White objects (see Figure 2.2) can be very different both on a macroscale and microscale level, like for example milk and porcelain. What they have in common is that they are made of an almost absorptionless material and inclusion of scatters. If the bulk of the material is large enough, the scattering process takes place to such an extend that all incoming light is scattered diffusely back. In clouds, water droplets acts as scatters; in paper, it is fibres; in white hair small air bubbles are present, and so forth.

Having an all-reflecting material that does not scatter light diffusely to some extent will therefore not be perceived as white. More likely, it will be perceived as mirror- or metal-like. Silver is a metal which reflects almost all visible light equally, see Figure 2.7. Even though it is perceived whiter than other metals, its appearance is very different from the distinctive anodised aluminium colours seen in Figure 1.1. It is clear that the optical mechanisms to produce white aluminium are fundamentally different than its colourful counterparts – and how to get around that is the core of the ODAAS project.

3 Rendering and visualising appearance

The true nature of most physical phenomena, including the propagation of light, becomes evident when simple elegant theories and mathematical models conform to experimental observations.

Harvey and Krywonos (2006)

The colour conversion equations in (2.1) are a key to modelling and quantifying visual appearance[1]. This chapter introduces the basic quantities used in colour and visual appearance analysis, including concepts from computer graphics.

3.1 Radiance

The basic quantity for all subsequent calculations is *radiance*. This is because the response of the eye is proportional to radiance (Dutré et al., 2006, p. 25) and its mathematical properties are in many cases more intuitive than for other quantities (Harvey et al., 1999; Harvey, 2012). This will become clear from calculations in the thesis and appended publications. Radiance is *not* to be confused with irradiance or intensity (Harvey, 2012). It is defined as the amount of power per unit solid angle per unit *projected* source area, which means that the angle between a surface and observation point is taken into account. Mathematically it is defined as

$$\text{Radiance} \equiv L = \frac{\partial^2 P}{\partial \omega_c \partial A_s \cos\theta_S}, \quad \text{(watts/steradian projected area)}, \tag{3.1}$$

where $\partial^2 P$ is the differential power flow through the differential solid angle $\partial\omega_c$ reflected by the projected differential area $\partial A_s \cos\theta_s$, where θ_s is the angle between observer and the normal to ∂A_s, see Figure 3.1.

For measurements, light incoming on a surface is normally given as irradiance:

$$\text{Irradiance} \equiv E = \frac{\partial P}{\partial A_s}, \quad \text{(watts/area)}. \tag{3.2}$$

[1]Interestingly, the ability to quantify human sensing in mathematical terms is so far unique for sound and light sensing. It is not possible for smelling, tasting or touching.

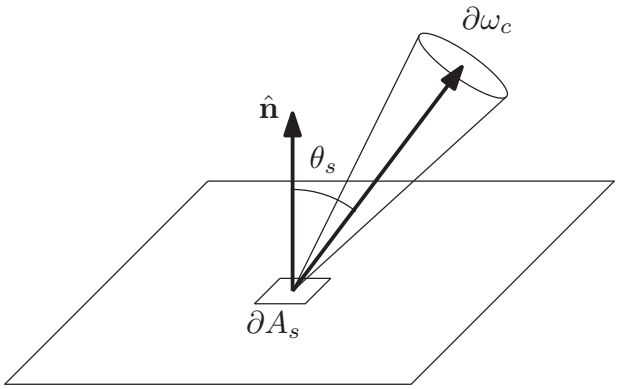


Figure 3.1: Geometry for the definition of radiance in (3.1).

Irradiance can be referred to as *intensity*[2], and it will be used in this thesis to denote irradiance upon surfaces as well as detectors. Converting irradiance to radiance is done through (Harvey et al., 1999)

$$L = \frac{\partial}{\partial\omega_c \cos\theta_s} E = \frac{\partial}{(\partial A_c \cos\theta_c / r^2)\cos\theta_s} E, \tag{3.3}$$

where θ_c is the angle with which the detector (collector) is deviating from being normal to the signal it is detecting, ∂A_c is the differential surface area of the detector, and r is the distance between ∂A_s and ∂A_c. Having a detector at a constant distance from the sample and always facing the sample, the following proportional relation is therefore valid

$$L \propto \frac{E}{\cos\theta_s}. \tag{3.4}$$

For all measurements performed in connection with the project, this is the case.

3.2 Coordinate systems

> *Often the simple nature of some natural phenomenon is obscured by applying a complex theory to model an inappropriate physical quantity in some cumbersome coordinate system or parameter space.*
>
> Harvey and Krywonos (2006)

Before continuing, definitions of coordinate systems used through the thesis are introduced. Taking outset in a traditional Cartesian coordinate system, spherical coordinates are de-

[2]The term may be defined differently in other sciences.

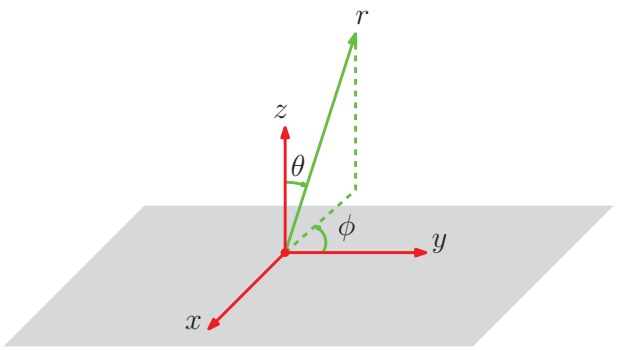


Figure 3.2: Relation between Cartesian and spherical coordinates. For the spherical coordinates $\theta \in [0, \pi[$, $\phi \in [0, 2\pi[$ and $r \in [0, \infty[$.

fined as in Figure 3.2. The space bounded by $\theta \in [0, \pi/2]$ is referred to as the *upper hemisphere*, and the space bounded by $\theta \in [\pi/2, \pi]$ is referred to as the *lower hemisphere*.

For different applications, a reformulation of the spherical coordinates is often beneficial (see the epigraph). With respect to scalar diffraction theory (SDT) (Harvey et al., 1999), it is beneficial to use the so-called direction cosine space defined as

$$\alpha = \sin\theta\cos\phi, \tag{3.5a}$$
$$\beta = \sin\theta\sin\phi, \tag{3.5b}$$
$$\gamma = \cos\theta. \tag{3.5c}$$

The coordinates make radiance shift invariant with respect to incident angle in SDT, which simplifies analytical calculations (Harvey et al., 1999). In direction cosine space, the hemispheres are bounded by a unit circle $\sqrt{\alpha^2 + \beta^2} \leq 1$, thus making them projected onto discs in the α, β plane. This makes it easy to visualise results.

For solving the radiative transfer equation (RTE) in Section 7.8 numerically, the coordinates

$$(\mu, \phi) = (\cos\theta, \phi), \tag{3.6}$$

will be used (see Appendix B). This coordinate pair simplifies RTE formulation and integration, but has the counter-intuitive property that $\mu = 1$ corresponds to $\theta = 0$ and $\mu = 0$ corresponds to $\theta = 90°$.

For other numerical implementations, it can be efficient to use

$$(\rho, \phi) = (\sin\theta, \phi), \tag{3.7}$$

since shift invariance for radiance can be conserved, and at the same time the discretisation is much more convenient than for direction cosines (P5). This system has the more intuitive property of $\rho = 0$ corresponding to $\theta = 0$ and $\rho = 1$ corresponding to $\theta = 90°$.

All spatial or angular coordinates used in the thesis refer to the above definitions.

3.3 The bidirectional reflectance distribution function

To optically describe surface reflections, the *bidirectional reflectance distribution function* (BRDF) is used. This function describes wavelength and angle dependent reflected radiance. Formally, for a certain wavelength λ, the BRDF is defined as

$$f(\theta_i, \phi_i, \theta_o, \phi_o, \lambda) = \frac{\partial L_o(\theta_o, \phi_o, \lambda)}{\partial E_i(\theta_i, \phi_i, \lambda)}, \tag{3.8}$$

where the subscripts i, o denote incident and outgoing directions respectively. The BRDF possesses several important properties which can be found in Dutré et al. (2006, p. 32). Most importantly for the thesis is that the normalized reflected energy can be calculated as

$$P(\theta_i, \phi_i, \lambda) = \int_0^{2\pi} \int_0^{\pi/2} f(\theta_i, \phi_i, \theta_o, \phi_o, \lambda) \cos\theta_o \mathrm{d}\theta_o \mathrm{d}\phi_o \leq 1, \tag{3.9}$$

where the inequality is due to energy conservation and the normalization of the BRDF.

BTDF, BSDF, BSSRDF and BSSTDF

A concept similar to the BRDF is the *bidirectional transmittance distribution function* (BTDF), which considers transmittance instead of reflectance. The combination of the two functions are sometimes described under one as the *bidirectional scattering distribution function* (BSDF). Many (semi-)transparent materials and interfaces are described adequately by a BSDF. Implicit in these expressions is the assumption that light is reflected/transmitted from the same point as it enters. This is adequate for all analysis in the thesis.

For completeness, it is mentioned that the more general *bidirectional surface scattering distribution function* (BSSRDF) is used when the point assumption fails. A classical example of the use of BSSRDF's is in the modelling of skin: In 2001 an Oscar award was given for constructing a BSSRDF (Jensen et al., 2001) that was used for modelling Gollum's skin in the film trilogy Lord of the Rings.

Lambertian diffuser

Ideally, materials like milk, clouds and paint reflect light uniformly over the entire hemisphere. This means that the reflected radiance for any given incident direction is constant for any given reflected direction. This property of perfect diffuse reflection is also called Lambertian reflectance, and the BRDF for such a surface is intuitively (Dutré et al., 2006)

$$f(\theta_i, \phi_i, \theta_o, \phi_o) = \frac{\rho_d}{\pi}, \tag{3.10}$$

where $\rho_d \in [0, 1]$ represents the fraction of incident light energy that is reflected. The upper limit of ρ_d is due to energy conservation, cf. (3.9). When a diffuse reflection or light intensity is mentioned in the thesis, it means that the *radiance* is (approximately) constant – not the intensity.

Specular reflectance

In contrast to the perfect diffuser one can imagine a perfect mirror, where incident angle equals reflection angle. This type of reflection is referred to as *specular reflection*. To conserve energy and at the same time restrict reflected radiance to one angle only, the BRDF is described as

$$f(\theta_i, \phi_i, \theta_o, \phi_o) = \rho_s \frac{\delta(\theta_i - \theta_o)\delta(\phi_i - (\phi_o + \pi))}{\cos\theta_i}, \tag{3.11}$$

where δ is Dirac's delta function[3] and $\rho_s \in [0, 1]$ represents the fraction of incident light energy that is reflected. In the thesis, specular reflection refers to light being reflected mainly around its *specular* reflection angle.

Glossy appearance

If the BRDF has an angular spread roughly centred at the specular reflection angle – instead of just a pure specular reflection – it is called *glossy*. Many different phenomena (like surface roughness) lay behind glossy surfaces, and there is no idealised BRDF describing this case. Different models such as the classical Blinn-Phong (Phong, 1975) or Cook-Torrance (Cook and Torrance, 1982) models are used within computer graphics for this purpose, but they do not necessarily represent strict physical analysis.

3.4 Computer graphics

All concepts in this chapter are part of basic computer graphics knowledge. In computer graphics, computational time is crucial, since rendering easily can become immensely time consuming[4]. For that reason, most BSDF models are normally approximations or heuristics which can be implemented efficiently. In several ways computer graphics is still relevant for the thesis and future work.

Simulating physical appearance

Already today, it is possible to use simulated data as input to computer graphics models. Such data are (very!) time consuming to generate, since structural colour models have to be simulated for a wide range of incident and reflected angles. This requires a 3D simulation model unless scattering in the third dimension is extracted from a 2D model. When the datasets are obtained, it is straight-forward to convert them into the same format as used in the MERL BRDF database (Matusik et al., 2003; MERL, 2014) which is a database of

[3]Note how this gives no upper limit on the amplitude of the reflected radiance.

[4]The 2011 Disney/Pixar film Cars 2 was rendered on a cluster using 12,500 CPU cores and having single frames taking up to 80 or 90 hours to render! `http://www.cnet.com/news/new-technology-revs-up-pixars-cars-2/`

BRDF measurements on 100 different materials. The data can be loaded in to Disney's BRDF Explorer (Disney, 2014) for quick visualisation and inspection. For computer graphics rendering, the freely available software Physically Based Rendering (`pbrt`) can be used (Pharr and Humphreys, 2010; pbrt, 2014). State-of-the-art for rendering structural colours based solely on physics is Okada et al., 2012 where a Morpho butterfly wing is simulated in 2D and then extracted to 3D and used for animations. In the computer graphics world, a render plugin called Maxwell Render (MaxwellRender, 2014) delivers a wide range of physics based computer graphics implementations.

All in all, the road seems paved for rendering photo-realistic scenes based on structural colour simulations. These can then be compared with sample photographies taken under similar condition, and hence lead to virtual prototyping of structural colour effects. This is an important step for visualising novel surface appearances like the results in P2 before selecting which one to fabricate. Furthermore, it will be a help for designers when fabrication technology catches up with all the novel designs presented in the thesis and many other places. Future work will hopefully accomplish this.

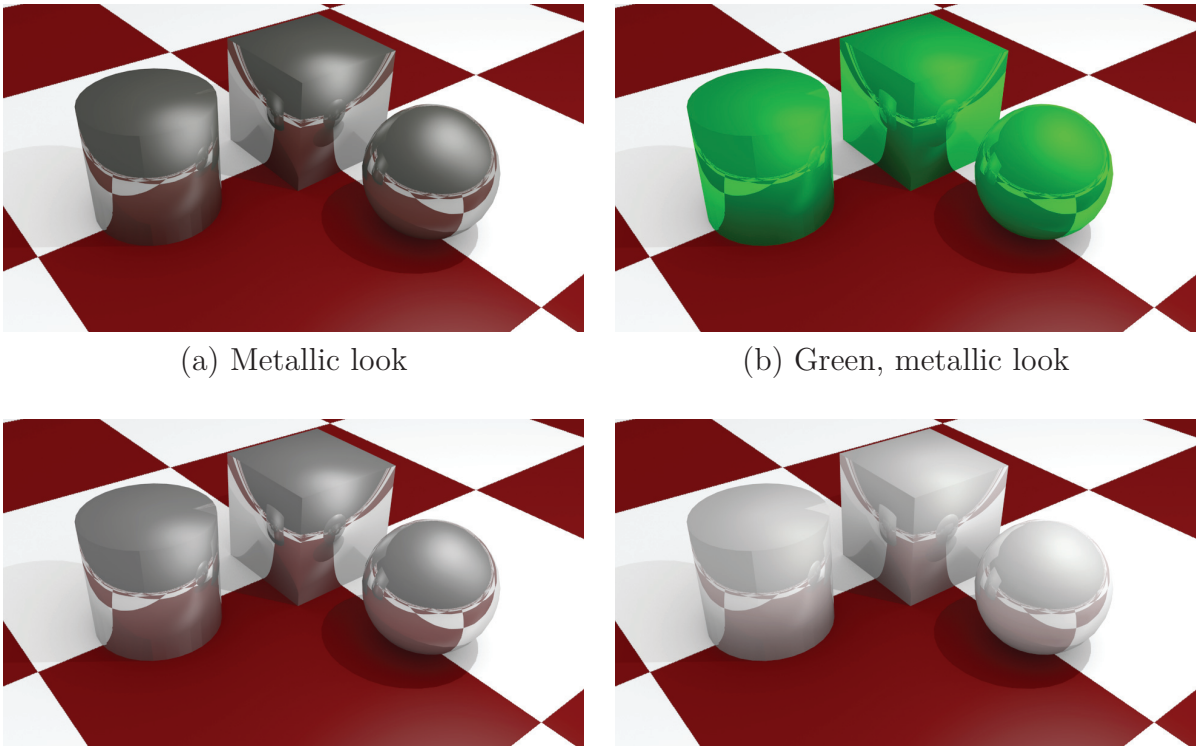
Rendering white aluminium

Even without physical models, computer graphics is used in this project to predict the desired appearance of white aluminium. This is to have an idea of what scattering properties ODAAS is aiming for in order to get a milky white, metallic appearance. Using the open source software Blender (Blender3D, 2014) to render objects based on a standard BRDF model with specular and diffuse/glossy reflection, the results in Figure 3.3 are produced. The whitish metallic appearance in Figure 3.3(d) is accomplished by adjusting the specular and glossy components to right proportions. From this it is learned that a strong diffuse scattering (much higher than what was expected before the experiment) is needed in order to get a whitish appearance. Unfortunately, there is no guarantee for energy conservation when fiddling with the different BRDF models in Blender. The result does therefore not predict whether or not white aluminium can be produced.

3.5 Photographing and visualising samples

As discussed in Chapter 1, the overall visual appearance of aluminium depends on alloy type and processing steps in an intricate way. Even though many effects can be captured using analysis, measurements and visual rendering, some effects are better captured and understood by visual comparison.

To convey the information, a photo setup was developed with emphasis on communicating the visual characteristics experienced in metals (strong specular reflection, and diffuse scattering for rough surfaces). The main problem is, that a highly reflective surface reflects its background, thus making the sample appear dark if the background is dark and bright if the background is bright. See an example in Figure 3.6(a). To accommodate this, a commonly known object (a die) was used on a dark background such that an ob-



(a) Metallic look (b) Green, metallic look

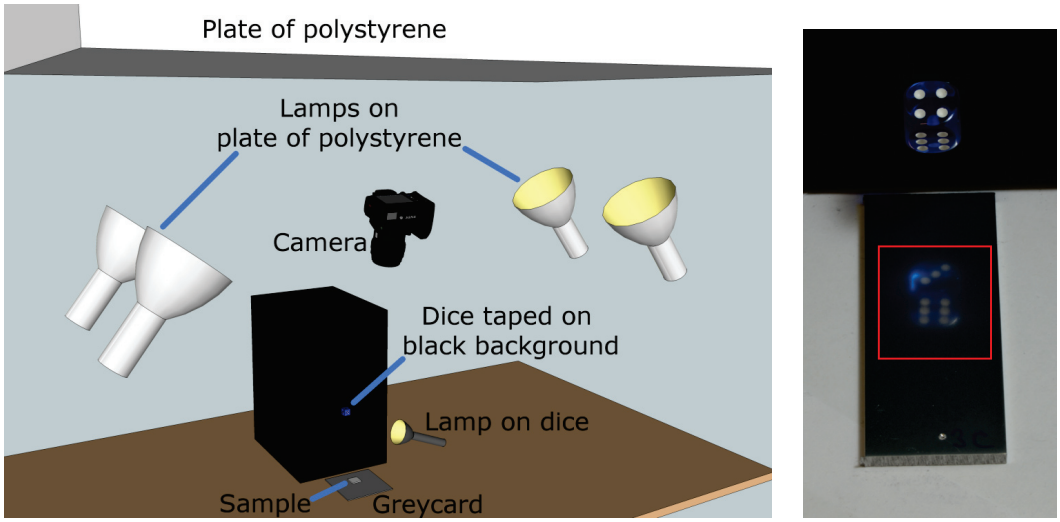
(c) Mirror-like, metallic look (d) White, metallic look

Figure 3.3: Rendering of different metallic appearances in Blender 3D. (a) A highly specular reflectance with some colour adjustments to imitate metal. (b) A highly specular reflectance of only the green colour component to simulate dyed metal. (c) A highly specular reflectance of all colours equally. (d) A specular reflectance of all colours equally plus some diffuse white scattering. Since these are computer renderings, there is no guarantee that the reflected energy does not exceed 100% – especially not for case (d).

server intuitively will understand that the background is dark, and sample brightness will be caused by diffuse lighting, whereas the specular properties are given by the replication quality of the die. Even small changes in purity of the specular reflection in aluminium are then easily detected. The setup is depicted in Figure 3.4(a) and has been used internally and some results are published in Aggerbeck et al. (2014a). All images are taken using the same camera settings to make the images comparable. Figure 3.4(b) shows an example of an image taken using this process.

As an example of its use, three photographs of aluminium anodised at different voltages are shown in Figure 3.5. It is seen how a low anodisation voltage gives the most pure reflection. Small differences like these are hard to capture in BRDF measurements or even by side-by-side comparison[5]. Comparing the images with earlier visualisations in the literature, this method improves the detail level of the representation of the appearances,

[5]Similar to Figure 2.4, even small changes in environment affects appearance significantly.



(a) (b)

Figure 3.4: (a) Photo setup used to capture the visual appearance of aluminium depending on treatment (Aggerbeck et al., 2014a). (b) Example of photography taken using the photo setup. The red square indicates the area which is cut out and used for comparison.

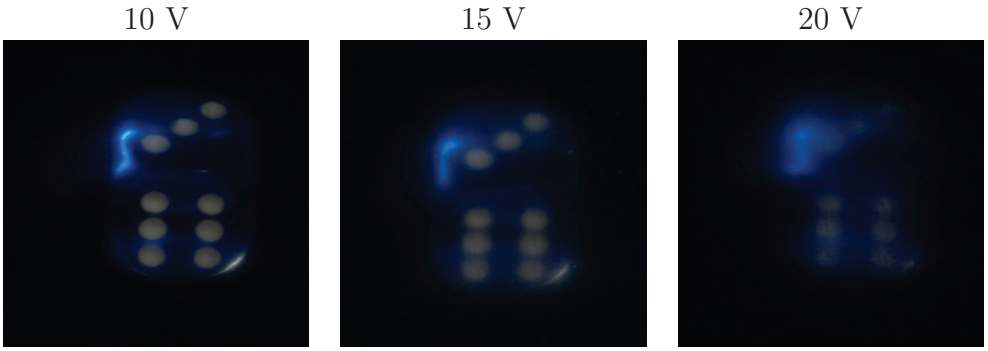
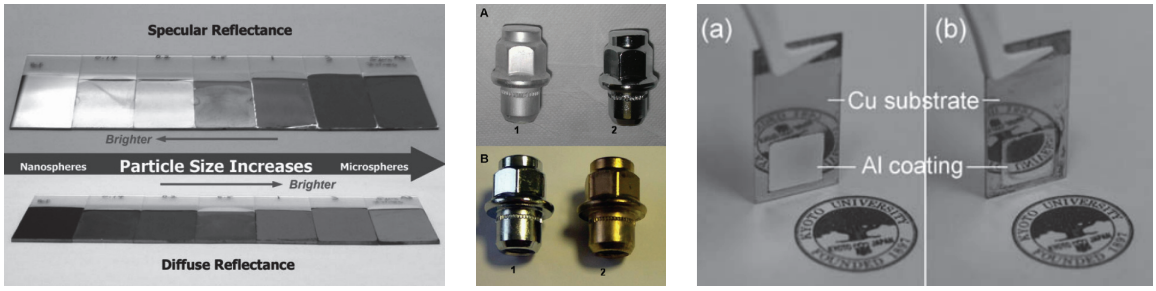


Figure 3.5: Photographs of polished Al6060 aluminium anodised to 10 μm using different anodisation voltages. Rather small differences in reflectance are readily seen by side-by-side comparison of the photographs.

see Figure 3.6. In the figure, the quality of the specular properties are hard to compare: for example is it hard to tell if A2 or B1 have similar properties in Figure 3.6(b), and if any mirror effect is present in some of the samples in Figure 3.6(a). Figure 3.6(c) is an elegant and illustrative way to show a clear difference, but it is hard to extend this set up to visualise smaller differences in appearance.

Photos have served as an important tool in the project for inspection and visualisation of sample properties and communication of results. In Chapter 7 more samples investigated using this method are presented.



(a) (b) (c)

Figure 3.6: Three examples from the literature of photographic reproductions of metallic appearance: (a) Comparison of specular and diffuse properties of nickel coated particle monolayers on polyelectrolyte multilayer substrates (Ahn et al., 2007). (b) Comparison of different electroplated coatings for anti-corrosion and decorative purposes (Barchi et al., 2010). (c) Comparison of electrodeposited Al-Zr coatings from baths with and without addition of $ZrCl_4$ (Shiomi et al., 2012). Images reprinted with permission.

4 Wave properties of light

You just have Maxwell's equations, some material properties and some boundary conditions, all classical stuff – what's new about that?

Bill Barnes, foreword to Maier (2010)

As explained in Chapter 2, many colour appearances are based on structures with dimensions comparable to the wavelengths in the visible spectrum. To model this requires wave analysis as described by Maxwell's equations. Giving an in-depth explanation of the consequences of those relations is beyond the scope of the thesis (and most undergraduate courses on electromagnetism as well). This chapter serves as a recapitulation of the phenomena that are relevant to the thesis.

4.1 Maxwell's equations

A particularly convenient form of Maxwell's equations is the so-called time-harmonic form, which considers (monochromatic) waves that do not change frequency when interacting with materials. This is true for all electromagnetic analyses in the thesis. Since the fields do not change frequency, they can conveniently be cast as $\boldsymbol{\mathcal{E}}(x, y, z, t) = \mathrm{Re}(\boldsymbol{E}(x, y, z)e^{i\omega t})$ for a solution at a certain frequency $f = \omega/(2\pi)$ where $\boldsymbol{\mathcal{E}}$ is a time dependent representation of the solution and $\boldsymbol{E}$ only is dependent on position. In a linear, isotropic and homogeneous medium, Maxwell's equations can then be stated as

$$\begin{aligned}
\nabla \times \boldsymbol{E} &= -i\omega\mu\boldsymbol{H}, && \text{(Faraday's law)}\\
\nabla \times \boldsymbol{H} &= \boldsymbol{J} + \sigma\boldsymbol{E} + i\omega\epsilon\boldsymbol{E}, && \text{(Ampère's circuital law)}\\
\nabla \cdot \boldsymbol{E} &= \rho/\epsilon, && \text{(Gauss's law)}\\
\nabla \cdot \boldsymbol{H} &= 0, && \text{(No isolated magnetic charge)}
\end{aligned}$$

where $\boldsymbol{E}$ [V/m] is the electric field, $\boldsymbol{H}$ [A/m] is the magnetic field, $\boldsymbol{J}$ [A/m^2] is the volume density of free current, σ [S/m] is the conductivity, ρ [C/m^3] is the volume density of free charges, μ [H/m] is the medium's *permeability* and ϵ [F/m] is the medium's *permittivity*.

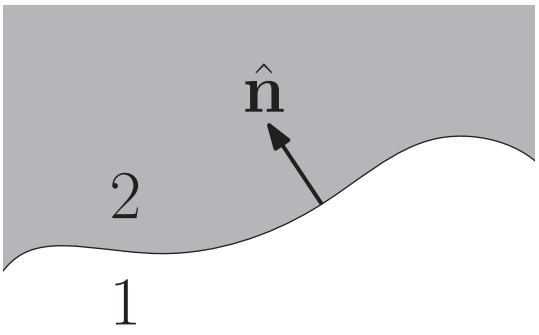


Figure 4.1: Geometry for boundary conditions between two media denoted 1 and 2. The normal vector is perpendicular to the surface at every point and pointing into medium 2.

Boundary conditions

To complete Maxwell's equations, boundary conditions between media are needed. Under the same assumptions as before, the boundary conditions between two media – as depicted in Figure 4.1 and subscripted 1 and 2 – are given by

$$\hat{\boldsymbol{n}} \times (\boldsymbol{E}_2 - \boldsymbol{E}_1) = \boldsymbol{0}, \tag{4.1a}$$
$$\hat{\boldsymbol{n}} \cdot (\epsilon_2 \boldsymbol{E}_2 - \epsilon_1 \boldsymbol{E}_1) = 0, \tag{4.1b}$$
$$\hat{\boldsymbol{n}} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{0}, \tag{4.1c}$$
$$\hat{\boldsymbol{n}} \cdot (\mu_2 \boldsymbol{H}_2 - \mu_1 \boldsymbol{H}_1) = 0. \tag{4.1d}$$

Implications of these boundary conditions are that the electric and magnetic field components *normal* to the interface are scaled by the ratio between permittivities or permeabilities respectively. Also, the *perpendicular* components to the surface are continuous across the boundary.

Fundamental parameters in optics

For most naturally occurring materials in optics $\mu = \mu_0$, where μ_0 is the free space permeability. This is implicitly assumed in the rest of the thesis. From the fundamental parameters of electromagnetism and the aforementioned assumption, some convenient parameters for optical analysis are defined as

$$\epsilon_r = \frac{\epsilon}{\epsilon_0} + i\frac{\sigma}{\omega\epsilon_0} \qquad \text{(Complex relative permittivity)}$$
$$n = \sqrt{\epsilon_r}, \qquad \text{(Complex refractive index)}$$
$$n' = \mathrm{Re}(n) = \sqrt{\frac{|\epsilon_r| + \mathrm{Re}(\epsilon_r)}{2}}, \qquad \text{(Real refractive index)}$$
$$n'' = \mathrm{Im}(n) = \sqrt{\frac{|\epsilon_r| - \mathrm{Re}(\epsilon_r)}{2}}, \qquad \text{(Absorption index)}$$
$$k = \frac{2\pi n}{\lambda_0}, \qquad \text{(Wave number)}$$
$$\boldsymbol{k} = k\hat{\boldsymbol{k}}, \qquad \text{(Wave vector)}$$

where ϵ_0 is referring to the permittivity of vacuum, λ_0 is the free space (vacuum) wavelength, and $\hat{\boldsymbol{k}}$ is a unit vector indicating a direction of a wave. In general *hatted vectors* will refer to normalised vectors. It it also common to denote the absorption index by k even though this can be confused by the wave number, but the context or a unit analysis can reveal which parameter is meant.

4.2 The wave equation

Maxwell's equations can conveniently be cast into a second order vector differential equation for either the electrical or the magnetic field. For the electric field it is expressed as[1]

$$\nabla \times \nabla \times \boldsymbol{E} - k_0^2 \epsilon \boldsymbol{E} = \boldsymbol{0}. \tag{4.2}$$

Even though the magnetic and electric fields are coupled, it can still be useful to recast the problem in terms of the magnetic field quantity. This recasting is e.g. used in P2 to simplify the numerical implementation.

Solutions in rectangular coordinates

Like other solutions to linear differential equations, the solutions to (4.2) can be stated as a (complex) weighted sum of a solution basis to the system. In a normal Cartesian coordinate system, where $\boldsymbol{k}$ is the complex wave vector and $\boldsymbol{x} = (x, y, z)$ is a point in space, the general vector solution to the wave equation is composed by complex vector superposition such that (Balanis, 2012, p. 105)

$$\boldsymbol{E} = \sum_j \boldsymbol{C}_j e^{\pm i \boldsymbol{k}_j \cdot \boldsymbol{x}}. \tag{4.3}$$

Using Euler's formula, the solution can for the lossless case ($\boldsymbol{k}$ is a real quantify) be decomposed such that the superposition is performed on the following components:

$$e^{\pm i \boldsymbol{k}_j \cdot \boldsymbol{x}}, \tag{4.4}$$

$$\begin{matrix} \cos(\pm k_{j,x} x), & \cos(\pm k_{j,y} y), & \cos(\pm k_{j,z} z), \\ \sin(\pm k_{j,x} x), & \sin(\pm k_{j,y} y), & \sin(\pm k_{j,z} z), \end{matrix} \tag{4.5}$$

where the exponential functions now only will be scaled to have information on travelling waves, and the sine functions describe the standing waves. The sign indicates the direction of the waves. For complex wave vectors, also evanescent waves and attenuating travelling and standing waves exist.

[1] Under the assumption that $\mu = \mu_0$.

4.3 Plane waves

Consider a wave propagating along the positive z-axis in an unbound, lossless medium. The wave vector will in this case be given as $\boldsymbol{k} = k(0\ 0\ 1)^T$ which means that the general solution for the electrical field can be constructed from (4.4) to take the form

$$\boldsymbol{E} = \begin{pmatrix} A_x e^{i\phi_x} e^{-ikz} \\ A_y e^{i\phi_y} e^{-ikz} \\ A_z e^{i\phi_z} e^{-ikz} \end{pmatrix}, \tag{4.6}$$

where A_x, A_y, A_z are amplitudes and ϕ_x, ϕ_y, ϕ_z are phases. Using Gauss' Law in this charge free medium yields

$$\nabla \cdot \boldsymbol{E} = 0 \quad \Leftrightarrow \quad -ikA_z e^{i\phi_z} e^{-ikz} = 0 \quad \Leftrightarrow \quad A_z = 0, \tag{4.7}$$

from which it can be concluded that $E_z = 0$. From Faraday's law, the magnetic field can be found as

$$\boldsymbol{H} = \frac{k}{-\omega\mu} \begin{pmatrix} E_y \\ -E_x \\ 0 \end{pmatrix}, \tag{4.8}$$

and it is therefore clear that only components perpendicular (transverse) to the propagation direction exist. Such waves are called transverse electromagnetic (TEM) waves or uniform plane waves. All electromagnetic waves that travel far in a (practically) unbounded medium will take the shape of (practically) uniform plane waves, and this makes plane waves an adequate description for light incident on a surface. All incident light waves are given as plane waves in the thesis.

4.4 Polarisation

For a plane wave described by (4.6) having $A_z = 0$, there are four degrees of freedom to define the wave by: A_x, A_y, ϕ_x, ϕ_y. Except the trivial choice of $A_x = A_y = 0$, tracing the electrical field vector over a full period draws an ellipse in the xy plane, see Figure 4.2. The shape, direction and orientation of this ellipse describes its *polarisation*. Particularly relevant cases of polarisation are *linear polarisation*, where the ellipse collapses to a line, and clockwise/counter clockwise *circular polarisation*. They are also depicted in Figure 4.2.

A rose by any other name would smell as sweet, and indeed, two linearly polarised, perpendicular fields can be referred to in many ways: E and H polarised; s and p polarised; or transverse magnetic (TM) and transverse electric (TE). The first designation in the mentioned pairs refers to the electrical field being perpendicular to a plane of incidence, whereas the latter refers to the electrical field being parallel to that plane of incidence.

The polarisation state of a plane wave might have significant influence on how it interacts with an object. For example, some 3D display technologies for TV and cinemas

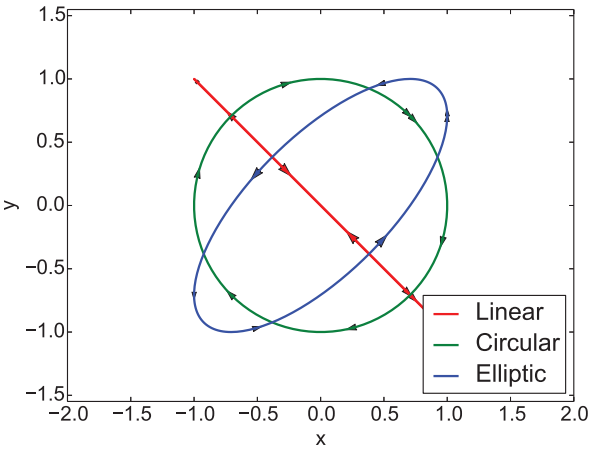


Figure 4.2: Traces of different polarisation states using the following parameters for $(A_x, A_y, \phi_x, \phi_y)$:
Linear $= (1, -1, \pi/4, \pi/4)$;
circular $= (1, -1, \pi/4, 5\pi/4)$;
elliptic $= (1, 1, 0, \pi/4)$.

are based on glasses with different polarisation filters for each eye. Photographers also use polarisation filters to enhance or minimise (specular) reflections, since, for example, reflections from plane surfaces are polarisation dependent.

4.5 Scattering from plane interfaces

For all surfaces that are well-approximated by a plane interface, their reflection and transmissions coefficients for TE and TM plane waves are described by the Fresnel equations

$$R_{TM} = \left| \frac{n_1 \cos\theta_1 - n_2 \cos\theta_2}{n_1 \cos\theta_1 + n_2 \cos\theta_2} \right|^2, \qquad T_{TM} = 1 - R_{TM}, \tag{4.9}$$

$$R_{TE} = \left| \frac{n_1 \cos\theta_2 - n_2 \cos\theta_1}{n_1 \cos\theta_2 + n_2 \cos\theta_1} \right|^2, \qquad T_{TE} = 1 - R_{TE}, \tag{4.10}$$

with a geometry as seen on Figure 4.3(a). The angle of the refracted/transmitted wave is given by Snell's equation as

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1}. \tag{4.11}$$

There are many implications of these results. It shows that larger contrast (difference between refractive indices) gives higher reflection; larger incident angles gives larger reflection as well; reflection is polarisation dependent; and it is possible to have 100% reflection above a critical angle if $n_1 > n_2$, this is called total internal reflection (TIF). At the same time it is also a good estimate of maximum obtainable reflectance for surfaces no matter their structure (Bao et al., 2002). Therefore, when structuring a surface like in P1, P4 or P5, this sets a limit which is unlikely to be surpassed.

As an example of Fresnel reflection and TIF, Figure 4.3(b) shows reflection (both polarisations) for an incident wave in air reflected by glass and vice versa. Here, TIF is observed, and also a significant difference in reflection between the two polarisations.

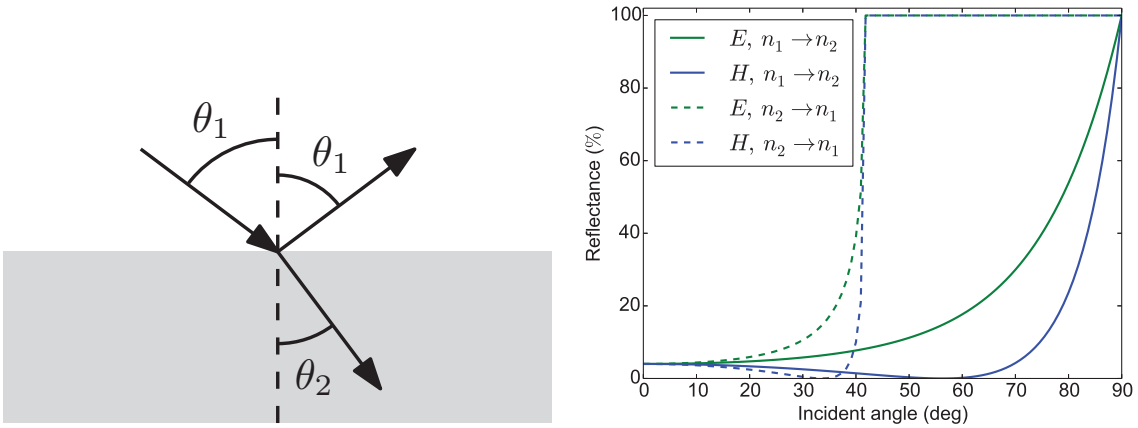


Figure 4.3: (a) Geometry for plane wave reflection. (b) Fresnel reflection for E and H polarisation for an air/glass ($n_1 = 1/n_2 = 1.5$) interface with the waves going either from n_1 to n_2 or from n_2 to n_1. Notice the total internal reflection for incident angles larger than $\sim 40°$ when going from glass into air.

Multilayer reflection

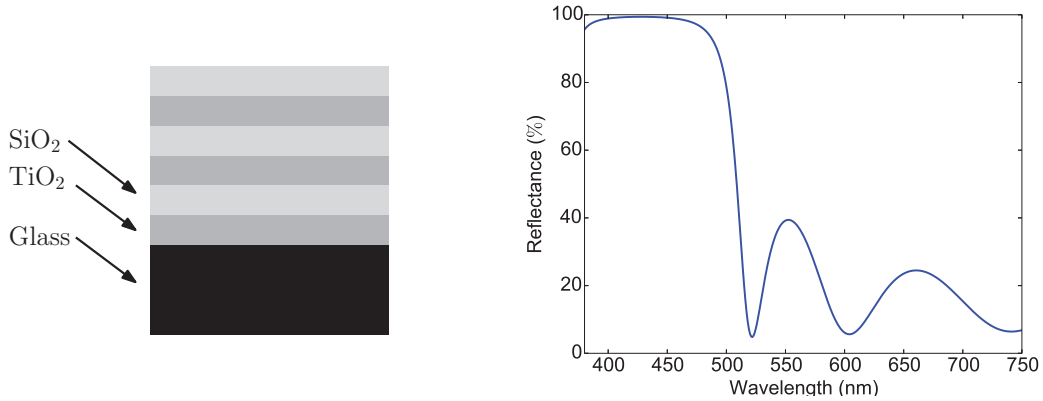
By stacking several plane media (see Figure 4.4(a)), it is possible to enhance reflection and transmission properties from that of a single layer (see Figure 4.4(b)). This is due to interference effects, and both full transmission and reflection can be obtained. The technology is used for applications such as optical filters and mirrors, anti-reflective surfaces, beam splitters and colour filters (see Figure 4.4(b)).

Assuming a solution of travelling (possibly attenuated) plane waves, complete solution of the multilayer system can be calculated by the so-called transfer matrix method (Yeh, 2005). The method is efficient and rather easy to implement, and a selection of free software to analyse multilayer systems are therefore available (Larouche and Martinu, 2008). Furthermore, several numerical methods exist for designing multilayer structures with a prescribed response (Tikhonravov and Trubetskov, 2012). This is discussed further in Section 5.2. Several examples of structural colour fabrication using multilayers can be found in the literature (Saito et al., 2009; Kolle et al., 2010; Chung et al., 2012; Kolle et al., 2013).

4.6 The far-field transformation

In practically all cases, coloured surfaces are observed optically far[2] away. Instead of simulating such long wave propagation[3], this can be predicted by using the Stratton-Chu equation (Stratton and Chu, 1939). This equation relates a scattered electromagnetic field

[2]Thousands of wavelengths.
[3]Which is completely infeasible!



(a) (b)

Figure 4.4: (a) Example of a multilayer structure. Using the right thicknesses, it corresponds to the structure used in Saito et al. (2009) to reproduce the blue colour of the *Morpho* butterfly. (b) Specular reflectance of the structure in (a) if 14 layers are used with TiO_2/SiO_2 thicknesses of 40/75 nm respectively (Saito et al., 2009).

$\boldsymbol{E}, \boldsymbol{H}$ on a closed surface S encompassing the scattering object by the following relation:

$$\boldsymbol{E}^f(\phi,\theta) = \frac{ik}{4\pi}\hat{\boldsymbol{r}}_0(\phi,\theta)\int\left[\hat{\boldsymbol{n}}\times\boldsymbol{E} - \eta\hat{\boldsymbol{r}}_0(\phi,\theta)\times(\hat{\boldsymbol{n}}\times\boldsymbol{H})\right]e^{ik\boldsymbol{r}\cdot\hat{\boldsymbol{r}}_0}\mathrm{d}S, \tag{4.12}$$

where $\boldsymbol{E}^f$ is the electrical far field in direction $\hat{\boldsymbol{r}}_0 = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, $\eta = \sqrt{\mu/\epsilon}$ is called the wave impedance and $\hat{\boldsymbol{n}}$ is the unit normal to S. This result means that only the electromagnetic interaction close to an object needs to be calculated, and from that, its far-field properties can be extracted. The principle is in different forms used in P2 and P3.

5 Optimised structural colour designs

Straight is the line of duty,
Curved is the line of beauty.
Follow the straight line, thou shalt see,
The curved line ever follows thee.

Isak Dinesen, The Monkey (Seven Gothic Tales)

As a first approach to understanding structural colour analysis, optimisation was used to investigate nanostructures with prescribed colour effects. The following sections present two different approaches: the first approach focuses on physical understanding and finding an optimisation formulation for structural colour problems. The latter focuses on designing complex structures based on the obtained experience. This is performed using full wave analysis, and it is shown how strong colour effects can be obtained.

5.1 Surface profile designs (P1)

To have an easy and fast framework for testing structural colour optimisation, surface profiles were simulated using scalar diffraction theory (SDT) which is a simple, approximate method for calculating reflection and transmission for interfaces. The simplicity of the formulation makes simulation extremely fast, and it is therefore well-suited for testing appropriate colour optimisation formulations.

Scalar diffraction theory

Consider a lit reflective surface profile, see Figure 5.1(a). According to Huygen-Fresnel's principle (Huygens, 1690; Balanis, 2012), *"every point which a luminous disturbance reaches becomes a source of a spherical wave; the sum of these secondary waves determines the form of the wave at any subsequent time"*[1]. This means, that when the incoming wave hits the surface, the reflection is found by summing all these spherical waves, each described by

$$E(R) = U_0 \frac{e^{-2\pi i R/\lambda}}{R}, \tag{5.1}$$

[1]The quote is from `http://en.wikipedia.org/wiki/Huygens-Fresnel_principle`.

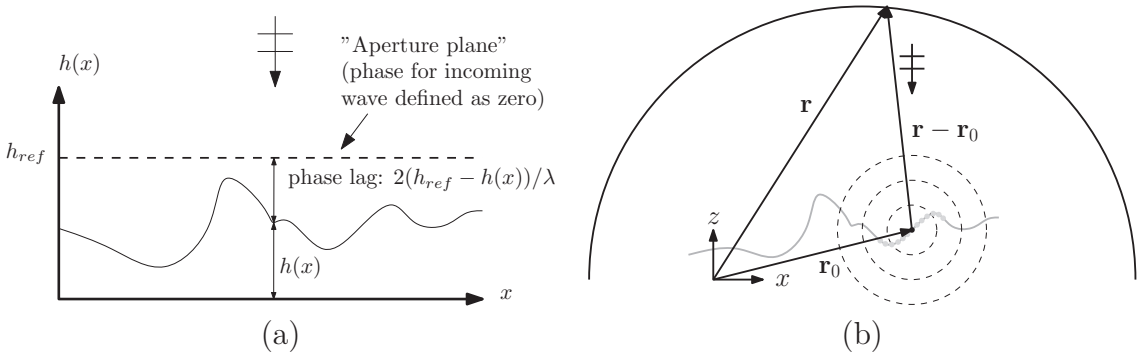


Figure 5.1: (a) Geometry for determining phase lag for a surface profile, when calculating its response using Huygen's principle. (b) Geometry defining the vectors used in far-field calculations when summing up spherical waves.

where R is the distance between source and observation point, and the complex number U_0 is the initial amplitude and phase of the spherical wave source.

Assuming that the spherical wave source is located in the point $\boldsymbol{r}_0$ and the observation point $\boldsymbol{r}$ is located far away (see Figure 5.1(b)) the distance to the object is found as

$$\begin{aligned} R &= \|\boldsymbol{r} - \boldsymbol{r}_0\| = \sqrt{\|r\|^2 - \|\boldsymbol{r}_0\|^2 - 2\boldsymbol{r}\cdot\boldsymbol{r}_0} \\ &= \|\boldsymbol{r}\|\sqrt{1 - \frac{2}{\|\boldsymbol{r}\|}\boldsymbol{r}_0\cdot\hat{\boldsymbol{r}} - \left(\frac{\|\boldsymbol{r}_0\|}{\|\boldsymbol{r}\|}\right)^2} \\ &\approx \|\boldsymbol{r}\|\sqrt{1 - \frac{2}{\|\boldsymbol{r}\|}\boldsymbol{r}_0\cdot\hat{\boldsymbol{r}}} \\ &\approx \|\boldsymbol{r}\|\left(1 - \frac{1}{\|\boldsymbol{r}\|}\boldsymbol{r}_0\cdot\hat{\boldsymbol{r}}\right) = \|\boldsymbol{r}\| - \hat{\boldsymbol{r}}\cdot\boldsymbol{r}_0, \end{aligned} \tag{5.2}$$

where the first approximation is due to ($\|\boldsymbol{r}\| \gg \|\boldsymbol{r}_0\|$) and the last approximation is due to the linear terms of the Taylor expansion $\sqrt{1-x} = 1 - x/2 - x^2/8 - \ldots$. This is the so-called phase approximation in the far field. When phase information is not needed, the approximation is further simplified to the amplitude approximation $R = \|\boldsymbol{r}\|$, since $\|\boldsymbol{r}\| \gg \hat{\boldsymbol{r}}\cdot\boldsymbol{r}_0$ as long as the origin is not located in the far field.

For a flat surface, the summation of spherical waves will again be a plane wave. Since the surface profile is perturbed, the phase of the spherical waves should also be perturbed accordingly – taking the extra phase lag of both the incoming wave and the point from where it is reflected into account. With reference to Figure 5.1(a), for a normal incident wave, this corresponds to

$$U_h(\boldsymbol{r}_0) = e^{-4\pi i h(\boldsymbol{r}_0)/\lambda} e^{4\pi i h_{ref}/\lambda}. \tag{5.3}$$

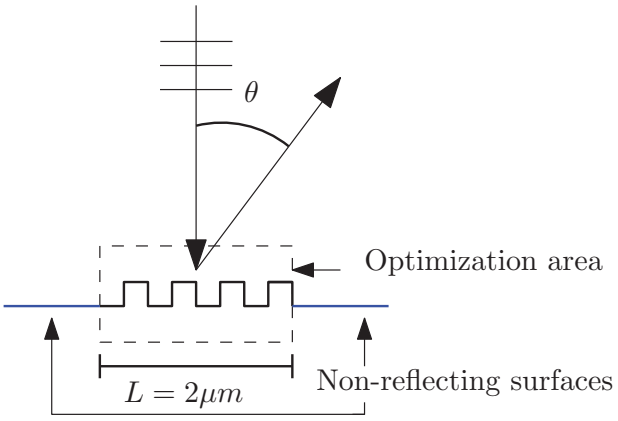


Figure 5.2: Geometry for the SDT colour optimisation problems: a 2 μm surface profile is modified in order to control colour reflection.

Summing all disturbed wave sources, the far-field reflected scalar field is then described by

$$\begin{aligned} E(\boldsymbol{r}) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U_h(\boldsymbol{r}_0)\frac{e^{-2\pi i(\|\boldsymbol{r}\|-\boldsymbol{r}\cdot\boldsymbol{r}_0)/\lambda}}{\|\boldsymbol{r}\|}\mathrm{d}x_0\mathrm{d}y_0 \\ &= \frac{e^{-2\pi i\|\boldsymbol{r}\|/\lambda}}{\|\boldsymbol{r}\|}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U_h(\boldsymbol{r}_0)e^{2\pi i\boldsymbol{r}\cdot\boldsymbol{r}_0/\lambda}\mathrm{d}x_0\mathrm{d}y_0, \end{aligned} \tag{5.4}$$

which corresponds to the Fourier transform of U_h multiplied by a constant phase correction and a factor $1/\|\boldsymbol{r}\|$ taking detector distance into account. By reformulating this with respect to direction cosines (Harvey et al., 1999, P1), the expression can be used to approximate reflections from structured surfaces. Even though this theory is simple (P1), it still serves as a good tool for understanding a wide range of phase induced scattering events, which is why it is used for initial investigation. Effects like surface resonances and plasmons are not captured by SDT, and a more thorough derivation of SDT can be found in e.g. O'Shea et al. (2003). Furthermore, for structures invariant in the y direction, SDT fits better with H polarised waves, since the electrical field vector of the incoming wave is then parallel to the structure, and the boundary condition is "less" violated by horizontal surfaces, which are not taken into account by SDT.

Optimisation for colour effects

Using the colour conversion described in Section 2.3, it is then possible to render the angle dependent colour reflection observed from reflections from different surface profiles. In P1, a 2 μm domain (see Figure 5.2) with normal incident light is used to investigate to what extend such profiles can be used to control colour appearance.

By discretising the surface profile into a vector of height bins $\boldsymbol{h}$, the structures are

designed by solving the following optimisation problem:

$$\begin{aligned} \max_{\boldsymbol{h}} \quad & f(\boldsymbol{h}) = \min_{i=1,2,\ldots,M} \frac{\|\boldsymbol{C}(\boldsymbol{h},\theta_i)\|^2}{\|\boldsymbol{C}_0(\theta_i)\|^2} - c_1 \frac{\|\nabla \boldsymbol{h}\|^2}{(Nd)^2}, \\ \text{subject to} \quad & \boldsymbol{h}_{min} \leq \boldsymbol{h} \leq \boldsymbol{h}_{max}, \\ & f_i(\boldsymbol{h}) = \frac{\|\boldsymbol{C}(\boldsymbol{h},\theta_i) \times \boldsymbol{C}_0(\theta_i)\|^2}{\epsilon^2 \|\boldsymbol{C}_0(\theta_i)\|^2} - 1 \leq 0, \end{aligned} \tag{5.5}$$

where $\boldsymbol{C}$ is the colour generated by $\boldsymbol{h}$, $\boldsymbol{C}_0$ is the prescribed colour, $\|\cdot\|$ is the Euclidean norm, f is the objective function, the f_i's are constraints on the reflected colour in a desired range, $\boldsymbol{h}_{min}$ and $\boldsymbol{h}_{max}$ are the extreme values that $\boldsymbol{h}$ can obtain, $\nabla \boldsymbol{h} = (h_1 - h_0, h_2 - h_1, \ldots, h_{N-1} - h_{N-2})$ is a measure of height differences between each element, N is the number of height bins, d is the width of each bin, c_1 is a parameter that penalises rapid changes in the surface profile, and ϵ is a parameter that sets the allowed derivation of an actual colour from its prescribed value.

The optimisation problem can be understood such, that given the colour constraints (f_i's) are fulfilled, then the angle with least relative intensity is maximised. This is seen by rewriting the fraction in the constraints such that

$$\frac{\|\boldsymbol{C} \times \boldsymbol{C}_0\|^2}{\epsilon^2 \|\boldsymbol{C}_0\|^2} = \frac{\left\|\boldsymbol{n} \sin\phi \|\boldsymbol{C}\| \|\boldsymbol{C}_0\|\right\|^2}{\epsilon^2 \|\boldsymbol{C}_0\|^2} = \frac{\|\boldsymbol{C}\|^2 \|\boldsymbol{C}_0\|^2 \sin^2\phi}{\epsilon^2 \|\boldsymbol{C}_0\|^2} = \left(\frac{\|\boldsymbol{C}\| \sin\phi}{\epsilon}\right)^2, \tag{5.6}$$

where ϕ is the angle between the two vectors. This means, that the constraints are fulfilled either by having a small angle between the vectors or a low energy reflection. The optimisation algorithm can therefore easily fulfil the constraints using low intensities and wrong colours, but in the process of maximising the energy, the colour deviation will also have to be minimised, thus going towards a desired solution to the problem. Details on the actual discretisation and implementation of the optimisation algorithm are found in P1.

Specular colour reflecting surface profiles

The structural colour optimisation method was performed for specular reflection of different colours for surfaces illuminated by normal incident light (P1). A selection of results are seen in Figure 5.3. Here it is seen how different colours require surface profiles of different kinds and different degrees of complexity. Blue and red can be obtained by fairly simple structures in 3 or 4 height levels, whereas green requires 6 levels. This is most likely because the red and blue colours each occupy their end of the visible light spectrum (see Figure 2.5(b)), making the structures act as low and high pass filters respectively. Green, on the contrary, is in the middle of the visible spectrum and a bandpass filtering is needed to filter out both the low and the high wavelengths of light. Thus requiring a more complex structure. Since red wavelengths are longer than blue (see Figure 2.5(b)), the structure for red is also expected to be of larger dimensions. It might seem surprising that the reflected

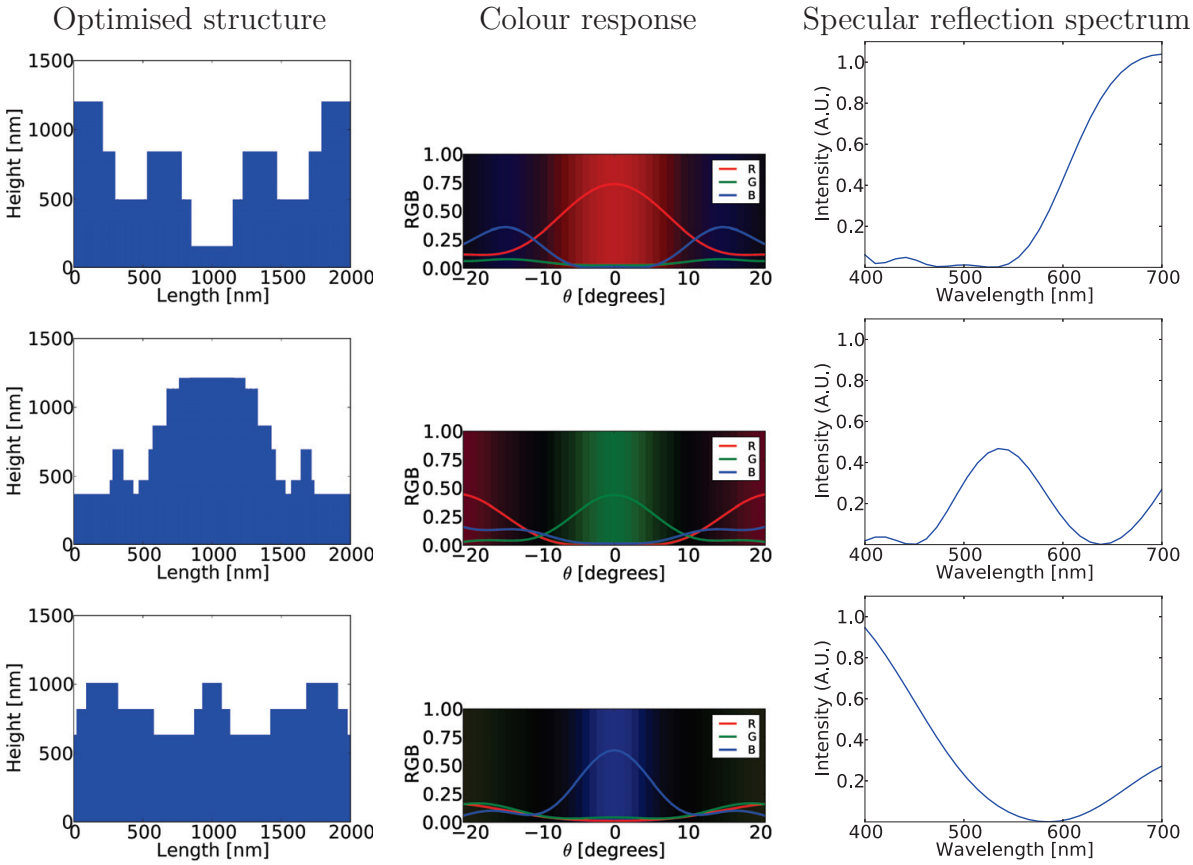


Figure 5.3: Results of specular reflected light optimisation for normal incidence for different colours using SDT (P1). The optimised structures shown in the left column have the colour responses shown in the middle column. The right column shows the reflected intensity spectrum at 0°.

blue spectrum has some red in it as well, but by considering the colour matching functions in Figure 2.5(a) this can be understood as a compensation for the negative red contribution at wavelengths around 450 nm to 550 nm.

The results verify the implementation of the optimisation, but due to the rather small finite size of the profile, it is difficult to get a more narrow angular reflection. This could in reality be obtained by either changing domain size or by considering a macroscopic distribution of the structures. For example, if the structures are distributed periodically with a period only allowing zero order diffraction, then the reflection will be truly specular. This is not possible with this formulation due to the limitations of SDT.

Wide angle colour reflecting surface profiles

Carrying out the same experiment for a flat angular colour response in the $\theta \in [-10°, 10°]$ range, the results in Figure 5.4 are produced. Here, it is again observed that blue gives a

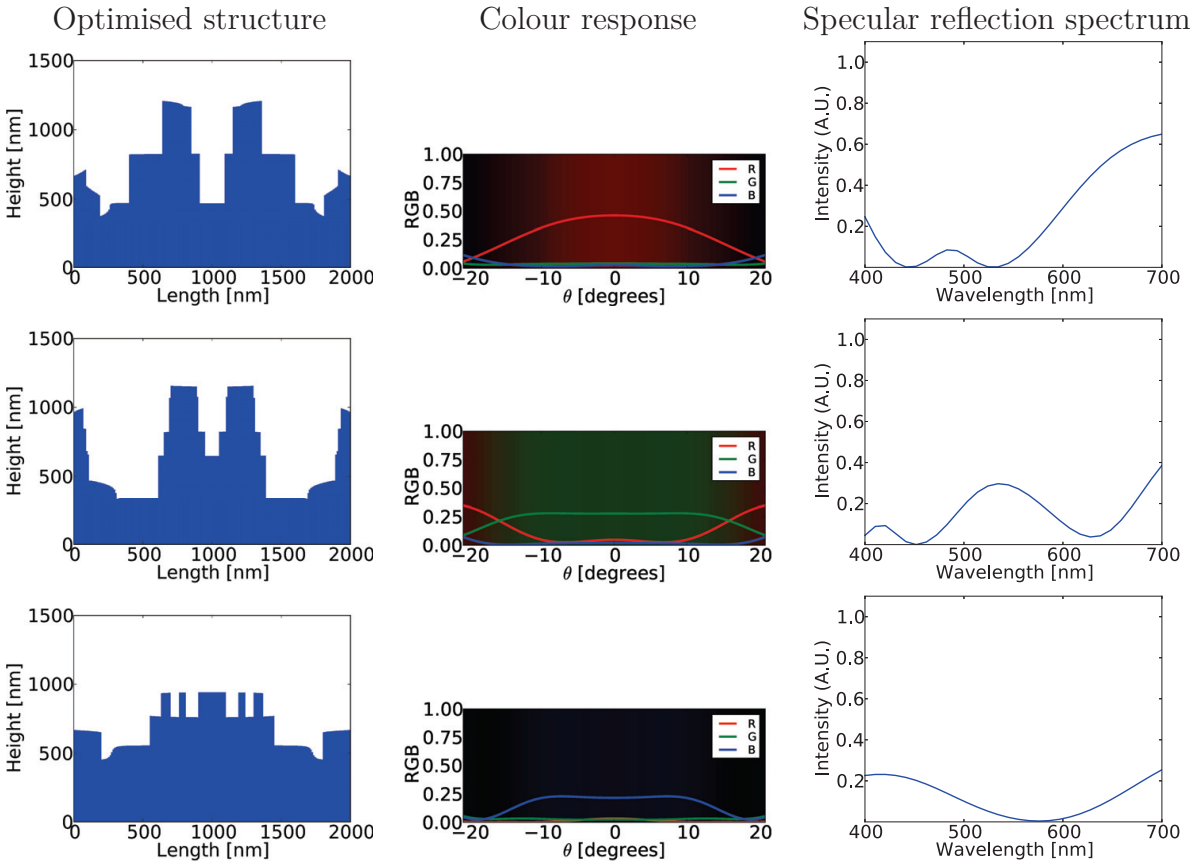


Figure 5.4: Results of specular reflected light optimisation for normal incidence for different colours using SDT (P1). The optimised structures shown in the left column have the colour responses shown in the middle column. The right column shows the reflected intensity spectrum at 0°.

simple design, whereas in particular green requires a more complex structure. It is difficult to obtain a flat red response, which is due to the fixed design domain size: since red wavelengths are longer, there is simply not enough room for the desired diffraction effect. To confirm this postulate, a 4 μm design domain was subsequently used for the red colour optimisation, and the result in Figure 5.5 shows that this did indeed solve the issue.

Full wave analysis of SDT results

In order to investigate the actual behaviour of the designs, a full wave analysis is performed (results were not included in P1). The results are shown in Figure 5.6, and a rough degree of correspondence is seen: the characteristic features from the SDT analysis are clearly present, but none of the designs exhibit the pure colours predicted by SDT. This was also not expected since the approximations are crude.

Even though the structures are not usable for design in their current state, the un-

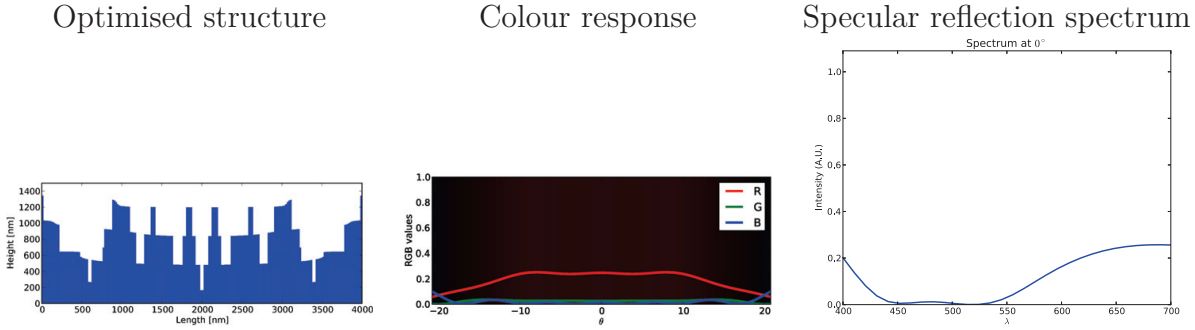


Figure 5.5: Results of a 4 μm surface optimised for red, wide angle light reflection using SDT. This result can be compared with Figure 5.4 where the same optimisation was performed for a 2 μm surface. Having a larger profile gives a more flat response.

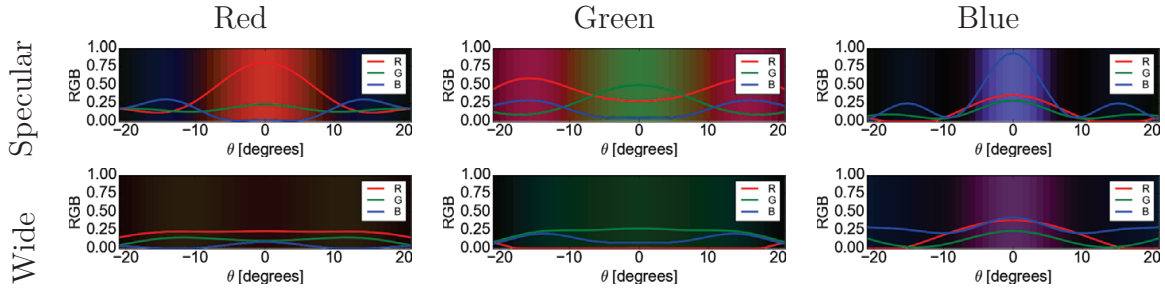


Figure 5.6: Full wave simulations of the designs in Figure 5.3 and 5.4 using the FEM model from P2. The incoming wave is H polarised, and since the code in P2 is not suited fur fully reflective surfaces, a constant dielectric material of $n = 2$ is used instead. Except for a proportionality factor, the change in response by varying n is minimal. The visualised results are all normalised using the same factor.

derstanding of basic design considerations, such as difficulties between different colours, dimension dependence of prescribed colour, the importance of domain size and how the optimisation algorithm performs is valuable for the analysis of structural colours. Furthermore, it shows that SDT is a good starting point for preliminary analysis and understanding of important phenomena within structural colouration. The method may also serve as generator of initial guesses for optimisation using the full wave method which will be presented in the following section.

5.2 2D two-material design (P2)

Optimisation using FEM is also performed in order to obtain results in better agreement with the physics (P2). At the same time, more complicated structures than just surface profiles can be designed, since each FEM element in a design domain is given a design freedom. This procedure is called topology optimisation (Bendsøe and Kikuchi, 1988; Bendsøe and Sigmund, 2004) and implementation details can be found in (P2) and Friis

and Sigmund (2012). Several methods for topology optimisation of wave propagation using FEM exist already (Aage et al., 2010; Jensen and Sigmund, 2011; Otomori et al., 2013), but the application to structural colouration is novel. In the future it might furthermore be interesting to investigate whether finite difference time domain (FDTD) optimisation is more beneficial for simulations in such a broad frequency range as the visible spectrum. FDTD is used for topology optimisation of wave problems as well (Hassan et al., 2014).

The present optimisation method takes outset in a design domain with a period of 1 μm and a height of 600 nm. The boundary conditions are periodic, but the far-field response is calculated on the basis of one unit cell. It is assumed that a random arrangement of the unit cells can ensure little interference between cells and therefore having an overall response of a surface similar to that of a single cell (P2, P3). Instead of having fully reflective material, a combination of two materials, SiO_2 and TiO_2, are chosen to yield structures which in principle can be fabricated. Some optimisations are carried out for one polarisation only, whereas others are carried out for unpolarised light which is an intensity mean of E and H polarisation. This better corresponds to the polarisation from most visible light sources.

Multilayer structures

As a first investigation, it is tested how well the approach can handle specular reflection if only horizontal (multi-)layers are allowed. This is in practice performed by locking all horizontal elements to each other in the optimisation algorithm. Some of the results from P2 are presented in Figure 5.7. Here, it is again seen how blue and red are easier to obtain than green. Furthermore, there is room for more details in the blue design since the domain is optically larger. The reason for these behaviours are expected to be the same as for the SDT optimisation. The most noticeable difference is that the set up allows transmission of undesired colours (instead of reflecting them off angle), and that several layers allow much stronger colour effects. Like with the SDT optimisation, the geometry restricts reflections from being more narrow.

In P2 results are also presented without the multilayer restriction, showing how much stronger intensities then can be obtained.

Result comparison with OpenFilters approach

The results in Figure 5.7 verify the feasibility of the approach, but it is a computationally expensive method for optimising multilayer structures. As discussed in Section 4.5, these systems can be solved in much simpler ways (Yeh, 2005) and due to the wide range of useful applications, there are well-developed methods for multilayer optimisation (Tikhonravov and Trubetskov, 2012). The optimisation algorithm in `OpenFilters` (*OpenFilters* 2014) repeatedly performs a gradient based thickness optimisation followed by a so-called needle method (Tikhonravov et al., 1996), where a new layer is inserted at an optimal position, until convergence of the needle method. Using this method, results of similar quality are obtained, see Figure 5.8.

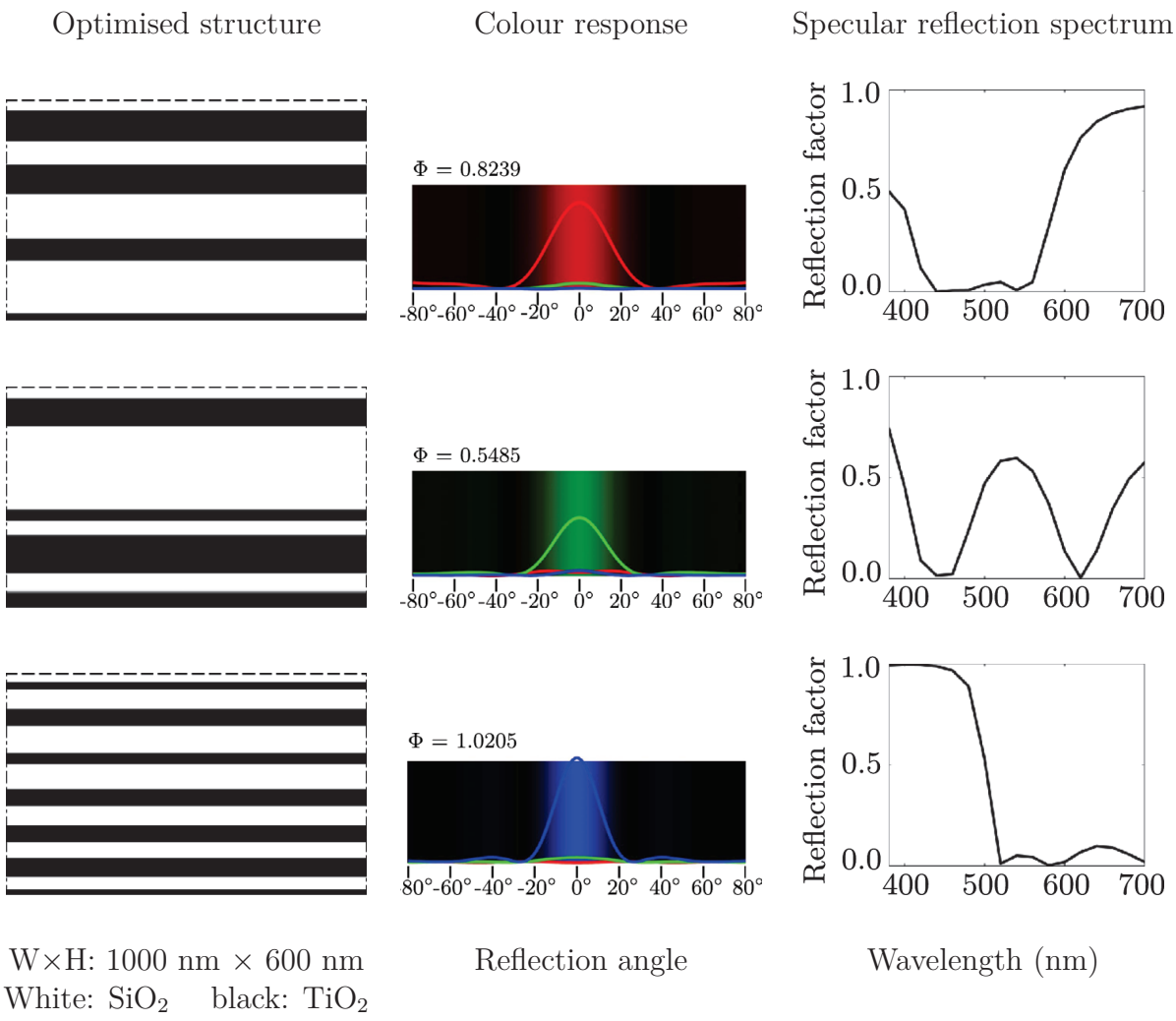


Figure 5.7: Multilayer optimisation for specular reflection of different colours using topology optimisation. The reflection factor is normalised such that the intensity at 0° for normal incident light on a fully reflecting surface corresponds to 1.

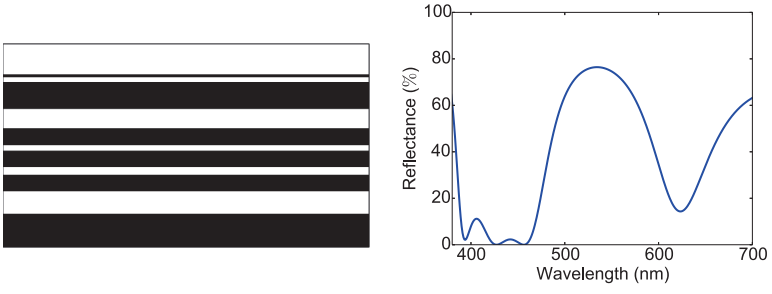


Figure 5.8: Design of comparable dimensions to Figure 5.7 for green specular reflection optimised using `OpenFilters` (*OpenFilters* 2014).

Within multilayer optimisation, a wide range of solutions can normally be found and the challenge instead lies in finding solutions that can be fabricated within given tolerances (Tikhonravov and Trubetskov, 2012). The method in P2 differs significantly from traditional approaches (Tikhonravov and Trubetskov, 2012), and implementing that approach in an effective multilayer solver (Yeh, 2005) combined with robust optimisation methods (Wang et al., 2011; Schevenels et al., 2011; Lazarov et al., 2012) might give designs more suited for fabrication. This is an interesting perspective for future applications – even though the field comprises more than 50 years of research in computational methods (Tikhonravov and Trubetskov, 2012), which probably leaves only few stones unturned. Future optimisation of multilayer structures can aim at finding solutions with low dependence on incidence angle, production tolerances or other features desirable for structural colouration.

Wide angle reflective structures

Relaxing the multilayer constraint and optimising for wide angle colour reflection ($\theta \in [-30°, 30°]$), the first three results in Figure 5.9 are obtained. First of all, it is noticed how the angular colour responses get wider and wider for increasing wavelengths (blue $\rightarrow$ green $\rightarrow$ red, see Figure 2.5(b)). This can again be related to the fixed domain width. Another feature of the wavelength dependence is that the smaller the wavelength of the prescribed colour is, the more detailed are the features in the design.

The colour selection mechanism in all three designs seems to be based on the principle of multilayer interference mixed with a reshaping of the layers to obtain reflection in the desired angular range. Compared to the previous results, all three colours are reflected almost equally strong – without a blue/red preference. The reason for this is not clear, but maybe the complexity required by the added constraints sets the limit for overall efficiency – instead of the "raw" colour selection process.

The possibility for more extraordinary colour effects are also investigated. A design optimised for red at $-43°$ and blue at $27°$ (the respective first order diffraction modes of the colours) for unpolarised light is presented as the last example in Figure 5.9. The result clearly shows that unusual, novel colour effects can be obtained as well. As of now, the results cannot be manufactured, but advances in nanoscale manufacturing (Ahn and Guo, 2009; Worgull et al., 2013; Röhrig et al., 2013; Takai et al., 2014; Christiansen et al., 2014) bring hope for future investigation of produced samples with novel colour effects.

Microwave scale model verification

Since the designs are difficult to verify experimentally, a microwave scale model experiment is planned instead. By extruding the designs and using additive manufacturing (3D printing), it is possible to design air/material structures with the same properties shifted down to the microwave frequency domain (10 to 26 GHz). The printing material (acrylonitrile butadiene styrene (ABS)) has been characterised by measuring the S parameters of printed samples inserted in a rectangular waveguide and then extracting the permittivity and per-

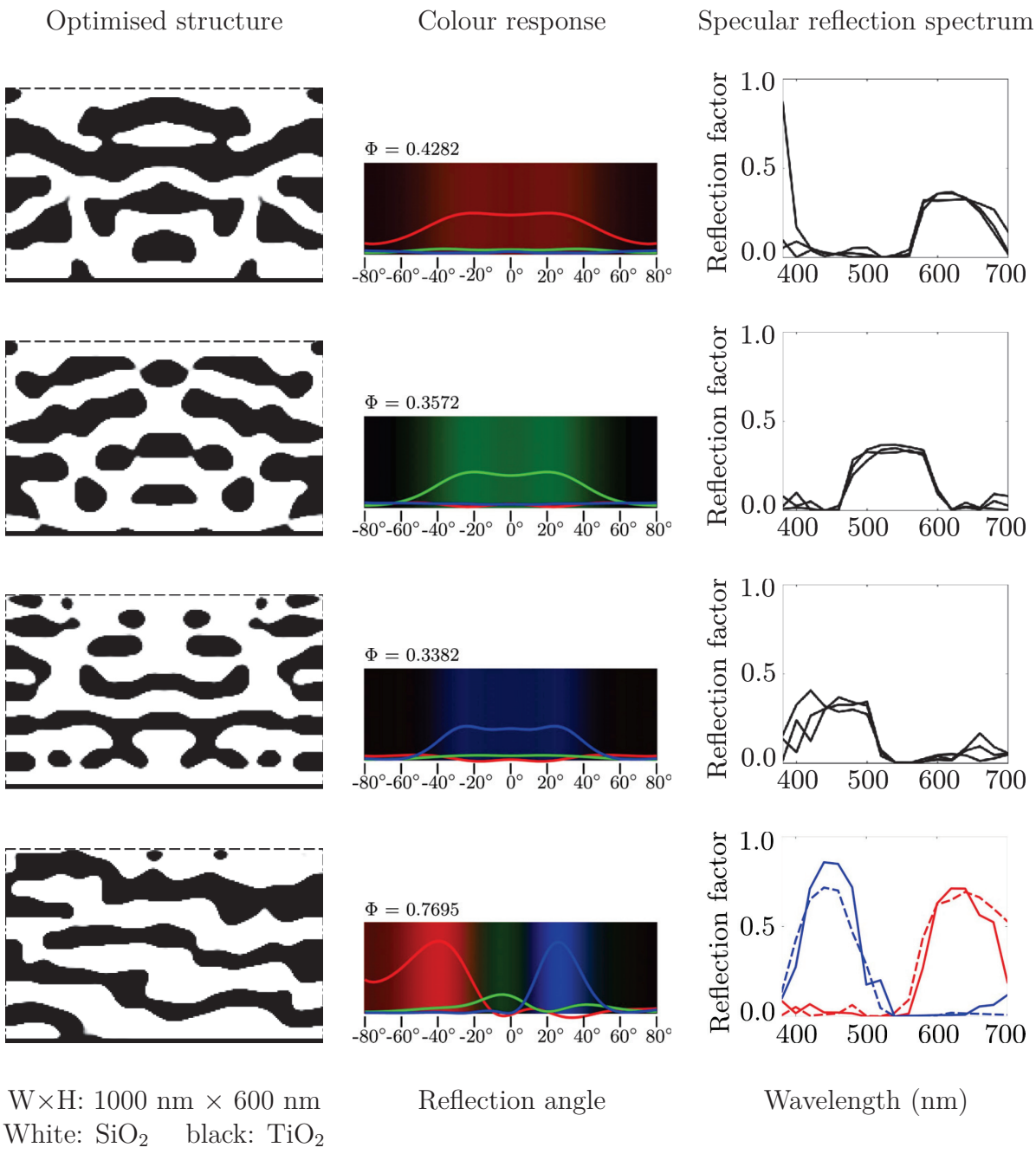
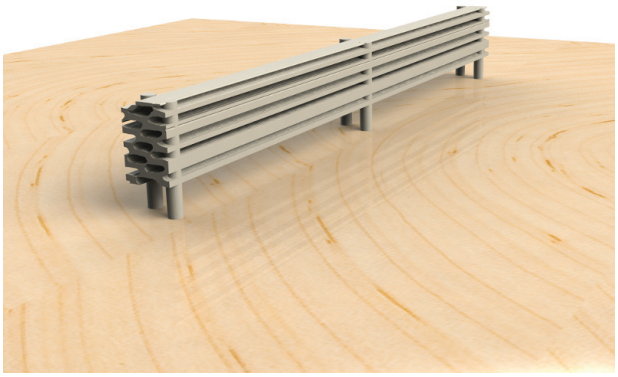


Figure 5.9: Topology optimisation with full design freedom. The first three optimisations are for wide angle ($-30°$ to $30°$) E polarised colour reflections, whereas the last optimisation is for red at $-43°$ and blue at $27°$ for unpolarised light. For the last optimisation, specular reflection for both E and H polarisation is seen. The reflection factor is normalised such that the intensity at $0°$ for normal incident light on a fully reflecting surface corresponds to 1.



(image: Erik Andreassen)

Figure 5.10: Rendering of a microwave scale model of a structural colour structure with the same blue/red scattering properties as the last design in Figure 5.9. The columns are added to support the structure and to ease mounting. Their influence is expected to be minimal. The width of the structure is 1.3 cm.

meability using the Nicolson-Ross-Weir method (Nicolson and Ross, 1970; Weir, 1974; Luukkonen et al., 2011). This subsequently requires re-running the optimisation with the updated material properties, and a different geometric design with the same downscaled scattering properties is obtained. A model of the design based on air/ABS for microwave frequencies is seen in Figure 5.10. The model is planned measured at University of Exeter where a set up is available (Vukusic et al., 2009).

This concludes the study of colour generating nanostructures. In P1 and P2 it is discussed how to distribute the obtained structures across a surface in order to avoid interference/diffraction effects. This will be discussed in the following chapter.

6 Order and disorder

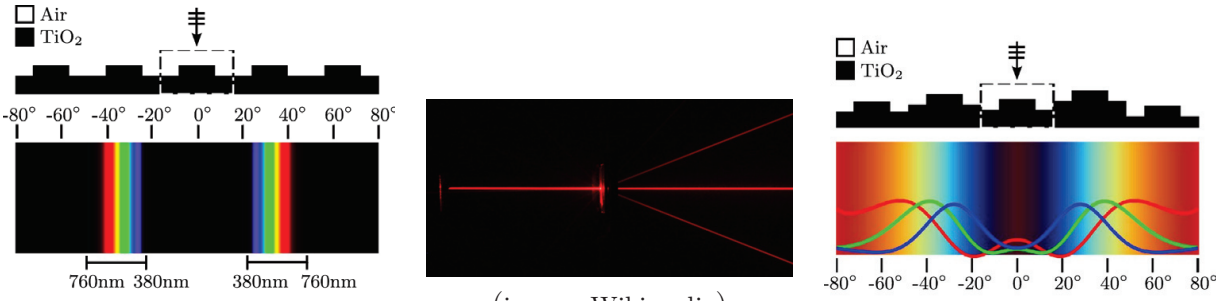
Nature has shown how little periodicity we can get away with.

Peter Vukusic, Ball (2005)

If the structures from the previous chapter are distributed regularly across a surface, it will give rise to diffraction patterns as seen in Figure 6.1(a), where each wavelength is reflected only in discrete angles given by the grating equation:

$$\sin\theta_i + \sin\theta_s = \frac{m\lambda}{d}, \qquad \text{(Grating equation)}$$

where θ_i is the incident angle, θ_s is the scattered angle, $m \in \mathbb{Z}$ is called the grating order, d is the period of the structure, and λ is the wavelength. For most people, this effect is highly unexpected the first time it is observed (see Figure 6.1(b)). The consequence of a structure regularly repeated across a surface is that the reflection properties are strongly angle dependent. This is not satisfactory for obtaining smooth colour effects, and other distribution methods (see Figure 6.1(c)) are investigated in this chapter in order to obtain desired appearances.



(a) (b) (c)

Figure 6.1: (a) The reflection from an incident plane wave on a regularly repeated structure (here with a period of 1 μm) will only reflect specific wavelengths at specific, discrete angles according to the grating equation. (b) Photo of a laser incident on a strictly periodic diffraction grating to illustrate the effect presented in (a). (c) If the periodicity is disturbed, by e.g. displacements in the height direction, the reflection behaviour is also changed, and more smooth variations are possible.

6.1 Nanorandomness (P3)

A best approach to estimate the total reflection properties of a single unit structure distributed across a surface is sketched in Figure 6.2. Here, the response from each instance is summed up in the far field, taking the phase difference between the different instances into account. This result is derived in P3 having the consequence that the total radiance of such a system can be found as

$$L(\theta,\phi) = \underbrace{\frac{r^2\eta}{2A_0\cos\theta}|\boldsymbol{H}_0^f(\theta,\phi)|^2}_{=\text{unit response}} \; N\underbrace{\left|\frac{1}{N}\sum_n e^{-ik(\hat{\boldsymbol{k}}-\hat{\boldsymbol{r}}(\theta,\phi))\cdot\Delta\boldsymbol{r}_n}\right|^2}_{=\text{Scaled Array Factor}}, \tag{6.1}$$

where the first factor is the calculated far field response for a single instance, and the last factor is denoted *scaled array factor* (SAF) (P3). In the SAF, $n = 0, 1, \ldots N-1$ is the instance index, $\Delta\boldsymbol{r}_n$ is the translation of the n'th instance, $\hat{\boldsymbol{r}} = (\sin\theta\cos\phi, \sin\theta, \sin\phi, \cos\theta)$, k is the wave number and $\hat{\boldsymbol{k}}$ is the normalised wave vector.

This approximation enables separate analysis of unit structures and their surface distribution, which makes it a powerful tool for analysis. A graphical interpretation of the concept is depicted in Figure 6.3. As described in P3, many observations in the literature (Lee and Smith, 2009; Saito et al., 2009, 2011; Steindorfer et al., 2012; Wang et al., 2013) can be explained from this analysis.

The grating equation

The Grating equation is easily proved using SAF's[1]. With reference to the coordinate system in Figure 3.2, having the plane of incidence as $\phi = \pi$, the incident light is described by the wave vector $\hat{\boldsymbol{k}} = (-\sin\theta_i,\, 0,\, -\cos\theta_i)$ and the element displacement by $\Delta\boldsymbol{r}_n = (nd, 0, 0)$ where d is the size/period of the unit structure. This gives rise to the SAF

$$\begin{aligned}\text{SAF}(\theta_s) &= \frac{1}{N}\sum_{n\in\mathbb{Z}} e^{-ik(\sin\theta_i-\sin\theta_s,\; 0,\; -\cos\theta_i-\cos\theta_s)\cdot(nd,\; 0,\; 0)} \\ &= \frac{1}{N}\sum_{n\in\mathbb{Z}} e^{-i2\pi nd/\lambda(\sin\theta_i-\sin\theta_s)}.\end{aligned} \tag{6.2}$$

The expression is identical to the so-called comb function (P3)

$$SAF(\theta_s) = \sum_{n\in\mathbb{Z}} \delta\left(\frac{d}{\lambda}(\sin\theta_i - \sin\theta_s) - n\right) := \text{comb}\left(\frac{d}{\lambda}(\sin\theta_i - \sin\theta_s)\right), \tag{6.3}$$

where δ is the Dirac delta function. The expression is only non-zero when

$$\frac{d}{\lambda}(\sin\theta_i - \sin\theta_s) = n \in \mathbb{Z}, \tag{6.4}$$

which is exactly the condition stated in the Grating equation. The comb function is sketched in Figure 6.3(b).

[1] The result is in this case exact, since no assumptions need to be made (P3).

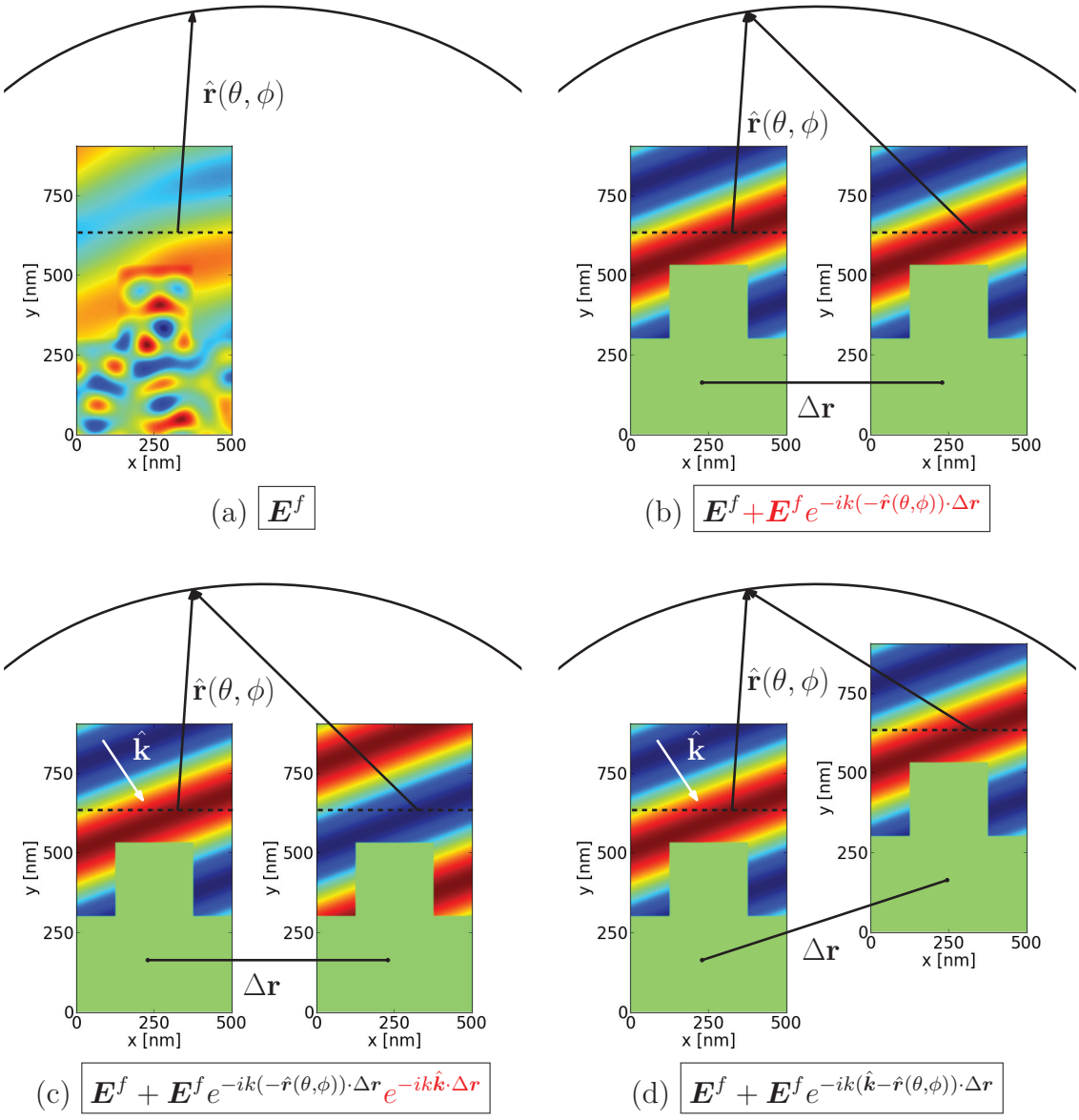


Figure 6.2: Graphical interpretation of the princicple behind deriving the scaled array factor. (a) The far field response is calculated for a structure based on its scattered field. (b) If a new instance is placed in the domain, its response relative to the first structure should be corrected for by a phase term due to the difference in distance to the observation point. (c) This phase term in itself is not enough, and an adjustment taking the shifted phase of the incoming plane wave into account is also needed. (d) By a simple dot product with the displacement vector, the phase difference can now be defined for arbitrary displacements.

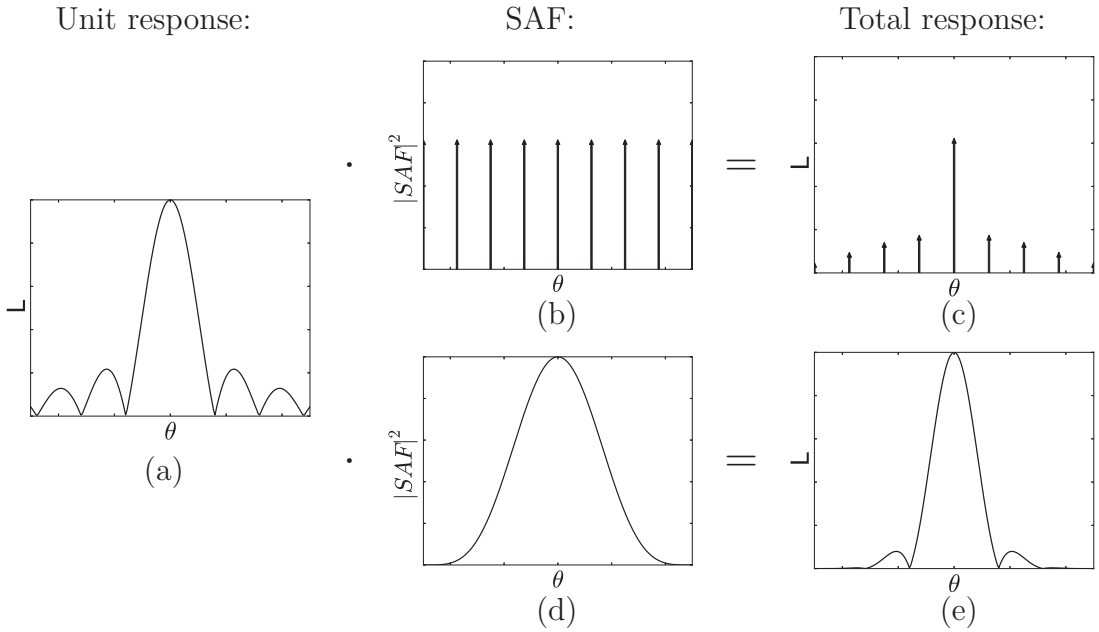


Figure 6.3: Example of how a structure with some unit response (a) is influenced by different SAF's. In (b) the SAF for a strictly periodic structure is shown (the comb function), and the effect of combining the structure giving the unit response in (a) with the pattern giving rise to the SAF in (b) is shown in (c). The arrows indicate that all energy is emitted at discrete points. If the SAF is like in (d), then the response will end up as seen in (e). (P3)

Similar results in the literature

The idea of separating phase modulation from structure response is not new. The concept is similar to the array factor (AF) for antenna theory and furthermore it can be related to the structure factor within condensed matter physics and the Debye-Waller effect (Licinio et al., 1999) as well. Also, the approach of summing phases has been used before for Morpho butterfly analysis (Kinoshita et al., 2002; Smith, 2009; Kinoshita, 2013; Yoshioka, 2013), but the results were based on SDT or point scatterers. The theories could therefore not be claimed to work for the structures being analysed, and therefore only worked as conceptual explanations. Articles are still produced that look into these randomisation effect (Steindorfer et al., 2012; Wang et al., 2013), and for this reason, the SAF was introduced as a tool for understanding the basic principles which are overlooked when only performing numerical analysis.

6.2 Coherence

In analysing light so far, it has been assumed that the light source – be it an incandescent light bulb, a fluorescent lamp, the Sun or a laser – can be modelled as a plane wave. This

assumption is ideal, since it does not take any temporal fluctuations of the plane wave into account.

Imaging that the surface of the Sun can be described by a vast amount of point sources – just like in SDT. The intensity observed on Earth from two such point sources (subscripted 1 and 2) can be described by

$$\begin{aligned} I &\propto \left|A_1 e^{i\phi_1} + A_2 e^{i\phi_2}\right|^2 \\ &= \left(A_1 e^{i\phi_1} + A_2 e^{i\phi_2}\right) \overline{\left(A_1 e^{i\phi_1} + A_2 e^{i\phi_2}\right)} \\ &= A_1^2 + A_2^2 + 2A_1 A_2 e^{i(\phi_1-\phi_2)} e^{-i(\phi_1-\phi_2)} \\ &= I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\phi_1 - \phi_2) \end{aligned}$$

where $\bar{\cdot}$ means complex conjugate, A_1, A_2 are the amplitudes of the two point sources, $I_1 = A_1^2, I_2 = A_2^2$, and ϕ_1, ϕ_2 are the phases of the received field from each source respectively. This means that the intensity can be anywhere in between $I_1 + I_2 - 2\sqrt{I_1 I_2}$ and $I_1 + I_2 + 2\sqrt{I_1 I_2}$, where the two extremes indicate negative and positive interference respectively. Considering N point sources, the same derivation would lead to

$$\begin{aligned} I \propto I_1 + I_2 + \ldots + I_n + 2\Big(&\sqrt{I_1 I_2}\cos(\phi_1 - \phi_2) + \\ &\sqrt{I_1 I_3}\cos(\phi_1 - \phi_3) + \ldots + \sqrt{I_{N-1} I_N}\cos(\phi_{N-1} - \phi_N)\Big), \end{aligned} \tag{6.5}$$

where the second summation contain all cross-correlations between phases. Due to random fluctuations of the source, the phases will vary over time. If this happens in an uncorrelated way, the time averages of the cosines will be zero, and the intensity will be proportional to the sum of the intensities from the individual sources. This state is called *incoherent*. In case the phase relation is fixed, the resulting wavefront is said to be *coherent*, and at the same time, the overall intensity from the same point source can potentially be much stronger. States in between – where some degree of correlation of the phases is present – are called *partial coherent*.

Going back to the example of the Sun: if the two point sources on the surface of the Sun are very close to each other, it is likely that their phases will be strongly correlated over time. On the contrary, if they are far from each other, this is highly unlikely. The degree to which they are correlated determines the coherence properties of the wavefront. The Sun and other normal light sources have a low degree of coherence, whereas a laser has an extremely high degree of coherence[2].

A general treatment of coherence is complex, and the explanation above will serve as an understanding of coherence in the thesis. For a deeper understanding, the comprehensive work from Mandel and Wolf is recommended (Mandel and Wolf, 1995), but in general most textbooks on optics have a short introduction to the subject (Hecht, 2001; Pedrotti[4], 2006).

[2]This also contributes to the laser's strong intensity since waves add constructively.

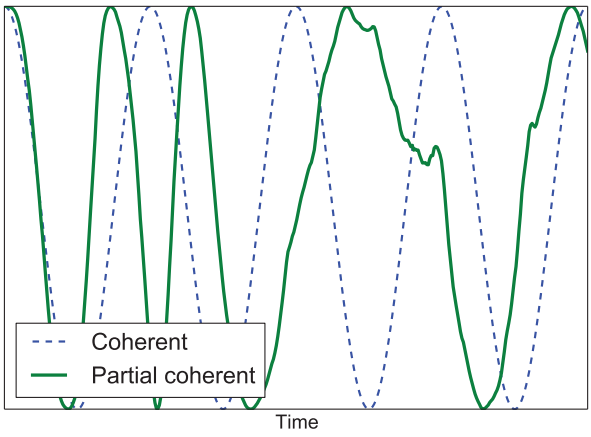


Figure 6.4: A partially coherent signal and a fully coherent signal for comparison. The further away from the origin, the less predictable is the actual phase based on knowledge from the ideal reference.

Impact of coherence on structural colour analysis

Coherence is normally divided into *temporal* and *spatial* coherence. The first describes the correlation between phases of light waves at different points along the direction of propagation, the latter describes the phase correlation between different points on a wavefront[3]. In the following, the impact of temporal and spatial coherence on the analysis in the thesis is considered.

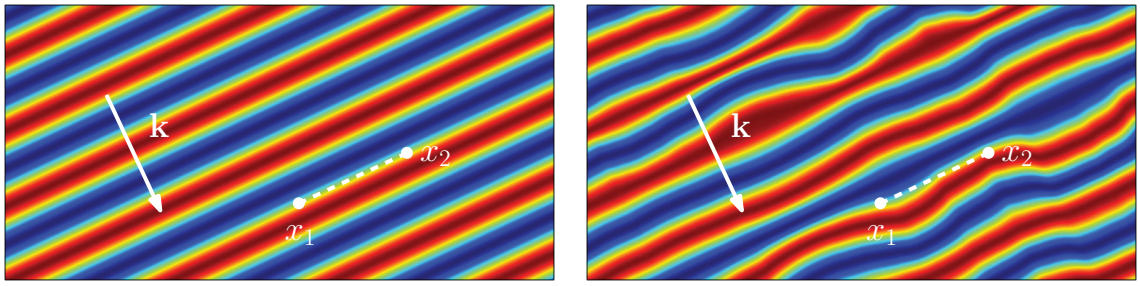
Temporal coherence

In case of low temporal coherence, interference effects (like multilayer interference) in the depth (z) direction will be weakened. This is because the wave cannot keep a constant phase relation with itself over time. A partial coherent signal with is seen in Figure 6.4. In practice, temporal coherence has not been taken into account in the project, since the structures based on interference effect in the thesis only are a couple of micrometers in extent. It has been visually observed during the project that interference effects in aluminium oxide drop dramatically around 7 μm. No structures presented in the thesis based on interference like effects get close to that limit.

Spatial coherence

As opposed to temporal coherence, spatial coherence cannot be neglected (Saito et al., 2011). The reason is, that whereas most structural colour structures extend some micrometers in the depth direction, they might easily cover a surface of centimetres or more in extent. Plane wave calculations over a large surface will therefore deviate from measurements with normal white light sources due to spatial coherence. An example of spatial coherence is seen in Figure 6.5. Measurements can be performed using lasers, which have an extremely high degree of coherence, but since the produced visual appearances should work for normal light sources, a way to model spatial coherence is needed to obtain usable results. This issue is treated in the next section.

[3]That is, perpendicular to the direction of propagation.



(a) (b)

Figure 6.5: Example of spatial coherence. (a) A fully coherent signal. By measuring one value (x_1), any other values (x_2) on the wavefront can be predicted. (b) A partially coherent signal. By measuring x_1, only values in the vicinity of that point can be estimated.

Correcting for coherence

For the moment, no obvious method exists for incorporating coherence properties of light in optical simulation models. Monte Carlo methods, long Finite Difference Time Domain (FDTD) simulations or methods tailored for specific geometries can be applied (Çapoğlu et al., 2011; Campa et al., 2013), but none of those methods are both efficient and general. In Saito et al. (2011), a reasonable trade-off is presented by averaging radiance from chunks of 50 units with a period of 400 nm, which each are simulated using coherent (plane) incident waves. This implies a total simulation domain of $50 \cdot 400$ nm $= 20$ μm. By varying the chunk size, the effect of varying spatial coherence can be estimated, but it is not possible to establish an exact relation between physical coherence properties and chunk size.

The approach in Saito et al. (2011) is well-suited for the proposed SAF based model in (6.1), since it does not require any additional simulations, but just averaging of the existing formula such that (P3)

$$|SAF(\theta)|^2 = \left\langle \left| \frac{1}{N} \sum_n e^{-ik(\hat{\boldsymbol{k}} - \hat{\boldsymbol{r}}(\theta,\phi)) \cdot \Delta \boldsymbol{r}_n} \right|^2 \right\rangle, \tag{6.6}$$

where $\langle \cdot \rangle$ means averaging over several instances of the sequence of randomised translations $\{\Delta \boldsymbol{r}_n\}_{n=0}^{N-1}$. This method has been used extensively throughout the thesis.

In general, incoherence has a smoothing effect on far field radiances. This is by no means strange, since coherence often can be expressed as a convolution in analytical calculations (Mandel and Wolf, 1995). During the project, work was initiated to investigate how this knowledge could be used on more complex systems. The results are shown in Appendix A, but were stopped since the averaging approach turned out to be adequate for the moment. The work can hopefully be continued, such that a method requiring no averaging and related to measurable sizes can be used in future work. This will make modelling faster and ensure that gradients of random parameters easier can be calculated analytically to ensure efficient, gradient based optimisation algorithms taking coherence into account.

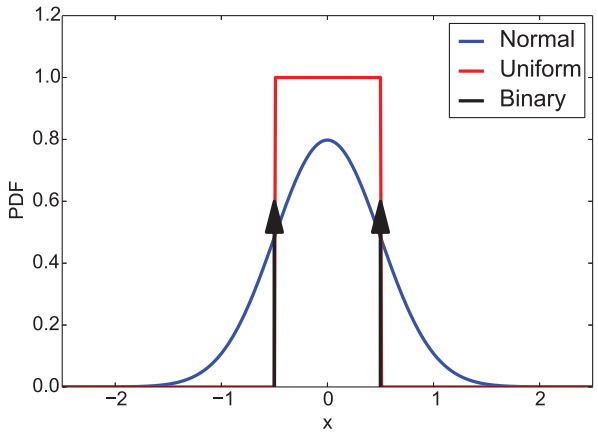


Figure 6.6: Three different PDF's used for designing randomisation. The PDF's are all centred around zero and with a maximum variation of $h_{max} = 1$. For the case of the normal distribution, this has been translated into a standard deviation of $\sigma = h_{max}/2$, since it has no fixed maximum.

6.3 Effect of randomisation

Using (6.6), different methods for distributing repeated structures can be investigated. It turns out that the most distinct changes arise due to change in the height (z) direction (P3). Even for this case, there is a remarkable difference in how this height distribution is chosen. To visualise this, the incoherent SAF of a 500 nm periodic structure has been investigated with $N = 40$ and 500 averages. The period for the SAF was chosen such that diffraction modes are present in a part of the spectrum for normal incident light (> 500 nm) in order to cover different types of behaviour, all in one simulation. Three common distributions are investigated: a binary distribution, a uniform distribution and a normal distribution. Their probability density functions (PDF) are shown in Figure 6.6.

For a height deviation parameter $h_{max} = 440$ nm, the corresponding SAF's can be seen in Figure 6.7. From here it is seen that grating modes are still strongly present for binary randomisation, weakly present for a uniform randomisation and not visible for randomisation based on a normal distribution.

A particularly desired SAF is the constant $|SAF|$, since this means that the unit response can be translated directly into the surface's reflectance properties. This desire is manifested in the optimisation articles (P1, P2). The deviation from ideal diffuse ($1/N$ everywhere) is calculated as

$$\text{deviation} = \frac{N|SAF|^2 - 1}{1} \cdot 100\%. \tag{6.7}$$

These deviations are calculated for different h_{max} and presented in Figure 6.8. As expected, binary randomisation does not seem to create a perfect diffuse SAF, but in a small angular range for a small range of wavelengths, a flat reflection can be obtained. Using a uniform distribution, a large h_{max} has to be chosen to suppress diffraction orders, and still they are visible in all simulations. Considering a normal distribution (this was not considered in P3) it is seen how even low choices of σ give a remarkably diffuse reflection! It is concluded, that the choice of random distribution of a unit structure across a surface is crucial to the overall colour appearance, and that a normal distribution seems favourable in order to obtain a constant SAF.

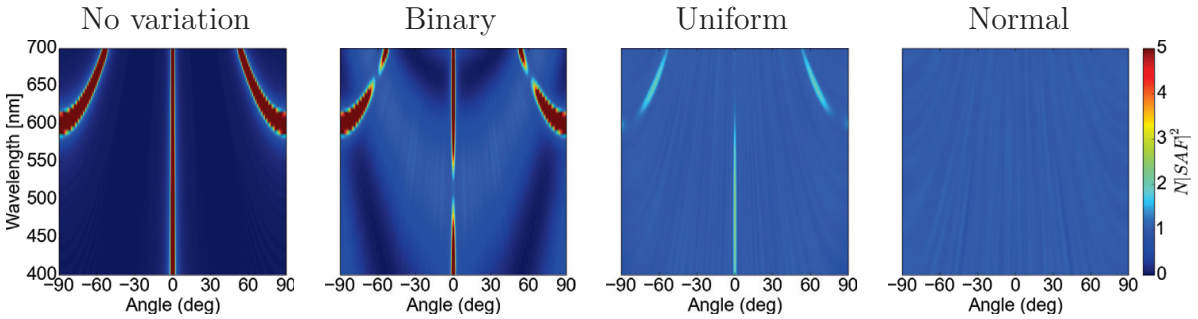


Figure 6.7: SAF for different distributions using $h_{max} = 440$ nm. A SAF with no randomisation is shown as well for reference. The grating modes here are clearly visible, and the little spread in intensity around the modes is caused by the finite length ($40 \cdot 500$ nm $= 20$ μm) of the structure.

Anisotropic randomisation (the *Morpho* effect)

The effects discussed so far have taken outset in 2D structures. An important consideration when producing actual samples is that isotropic reflectance in the full hemisphere might not be desired, and by concentrating the reflection in a smaller solid angle, the reflection will look more bright from most normal viewing angles.

This is exactly what is observed by Saito and stated as crucial for a proper reproduction of the *Morpho* butterfly's bright blue colour (Saito et al., 2004). The effect is best illustrated by Saito himself in Figure 6.9. This should be taken into account when bright, metallic colours – like white aluminium – are the desired design target.

6.4 Experimental verification (P4)

To verify the randomisation model experimentally, a silicon sample with a randomised diffraction grating has been produced using deep UV and dry etching (P4). A scanning electron microscope (SEM) image of the produced sample can be seen in Figure 6.10. The BRDF of the produced structure has been measured for normal incident light and compared to both the randomisation theory and a full wave simulation. The later is seen in Figure 6.11. A reasonable match was seen, and a detailed discussion of the results is found in P4. The most important findings from P4 are that the grating patterns caused by superperiods are much less pronounced in the measurements than in simulations; finer details like sharp resonances get more blurred; and large angle reflection predictions seem less accurate. Furthermore, the influence of SAF and unit structure can be identified from the measurements and simulation results (see Figure 6.11). This identification can help improving analysis of structurally coloured structures.

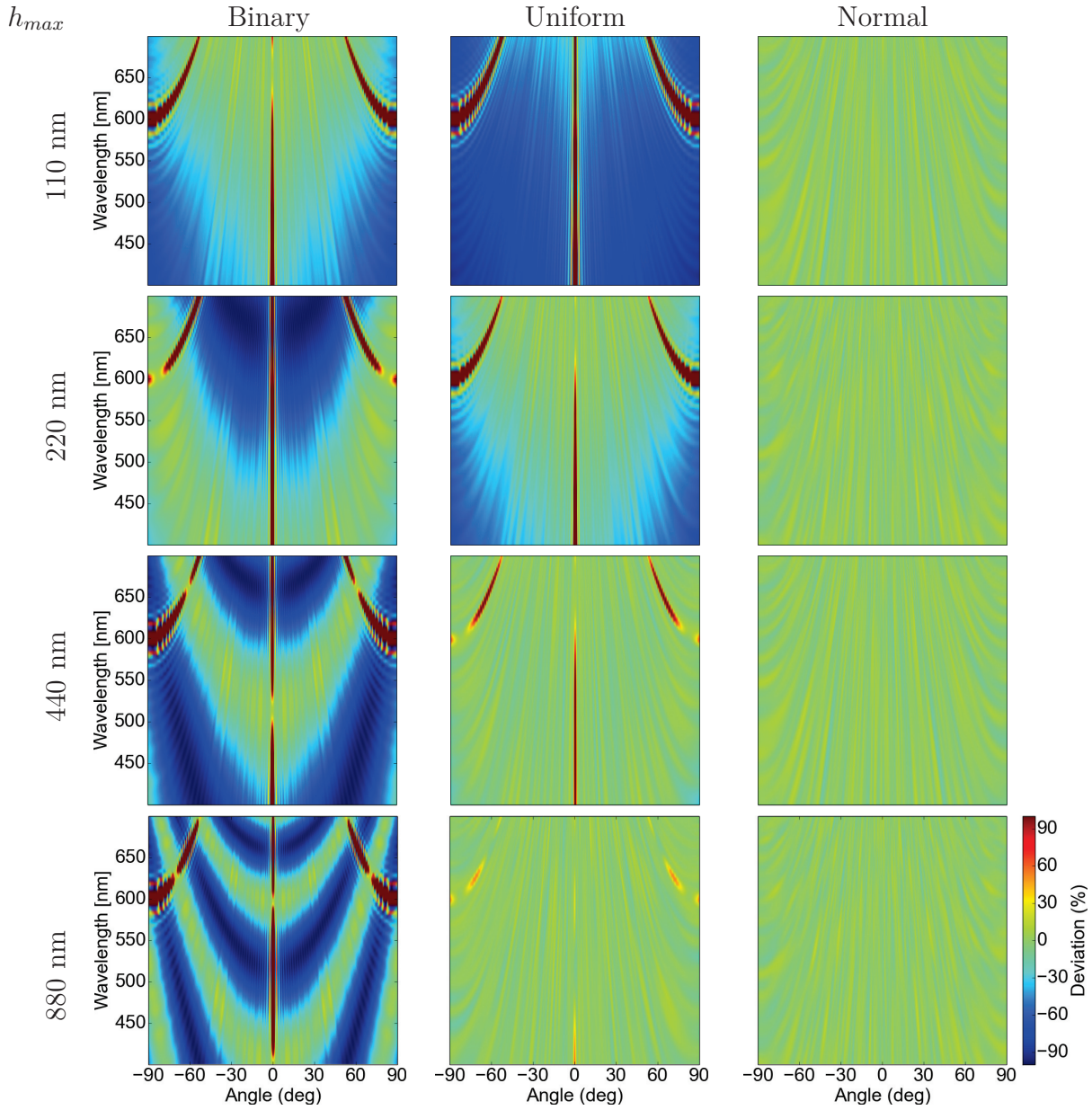


Figure 6.8: Deviation of SAF's from ideal diffuse at normal incidence calculated using (6.7). The results have been cropped to $[-100\%, 100\%]$ for a good colour representation. Even though 500 averages are used to incorporate coherence, superperiod diffraction modes are still visible. These are less pronounced in measurements (P4).

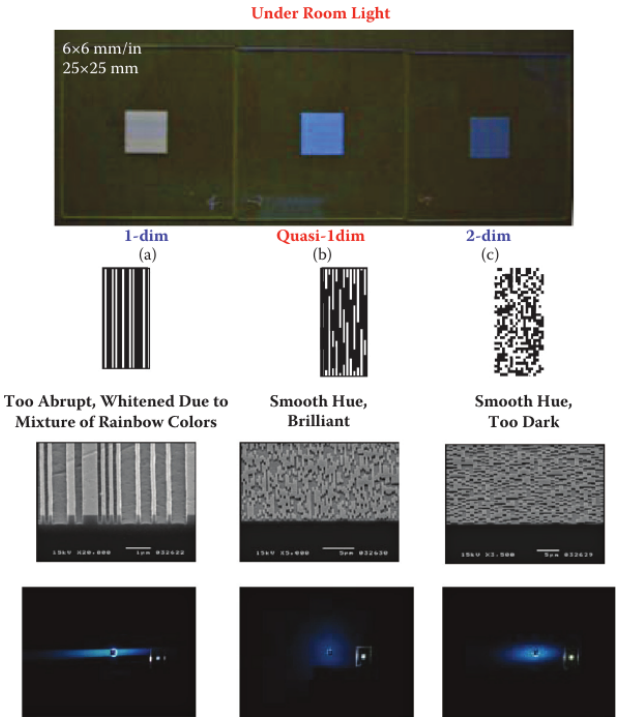


Figure 6.9: Figure 4.30 from a book chapter by Saito in Karthaus (2012). This exemplifies the effect of different types of randomness: the appearances observed under room light are multilayer coated versions of the randomised structures below. It is seen how an anisotropic randomisation gives a brighter blue than the more dull, isotropically randomised structure. The bottom row shows normal incident reflection onto a black screen (note here that the last two images to the right should be swapped).

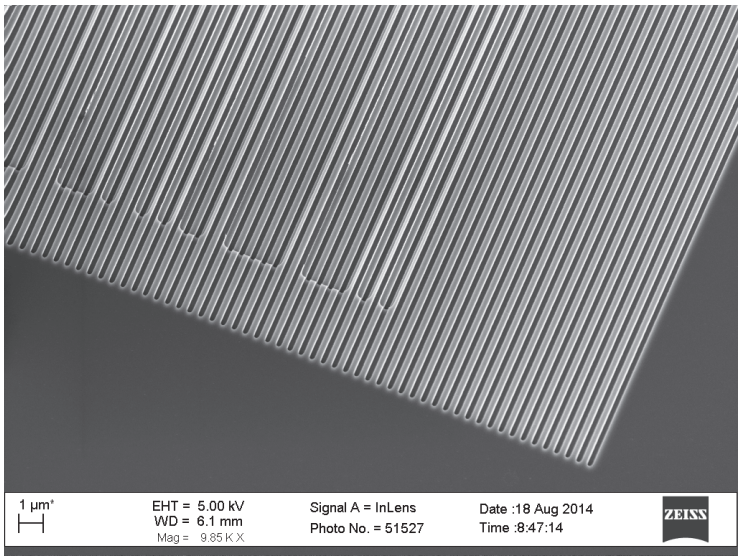


Figure 6.10: Corner of a silicon sample produced to experimentally verify the proposed randomisation theory: a normal line grating has been overlaid with random height variations.

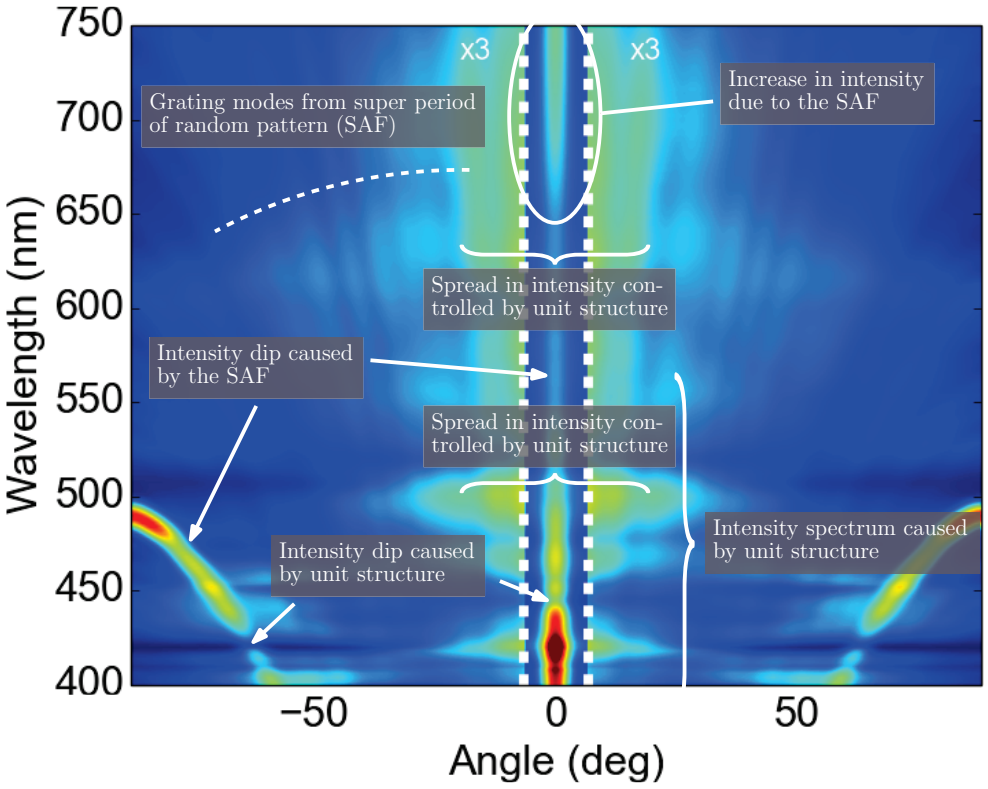


Figure 6.11: Full wave simulation of the incoherent reflectance spectrum for normal incident light of the structure in Figure 6.10. Several features and their origin have been identified on the plot.

6.5 Natural disorder

This chapter finishes the part on analysis based on deterministic geometries. The following two chapters are instead treating light interaction with stochastically described structures, with emphasis on modelling the optical properties relevant for anodised aluminium and white colouration.

Before continuing, it should be noted that even though the optical properties of the *Morpho* butterfly can be analysed using the presented approach (P3), many other kinds of disordered structures exist in nature which are far more complex to describe. Vukusic and co-workers have investigated several of these structures (Vukusic and Sambles, 2003; Vukusic et al., 2007; Vukusic et al., 2009; Pouya et al., 2011; Kientz et al., 2012; Burresi et al., 2014), and examples of such designs can be found in different beetle shells which give rise to extreme white, diffuse scattering (Burresi et al., 2014). See Figure 6.12. Examples of synthetically manufactured structures with a mix of "controlled disorder" and natural disorder also exist (Song et al., 2014).

By investigating such structures, it is clear that a whole world of opportunities within structural colours so far lies out of reach by only using "periodicity based" designs. To

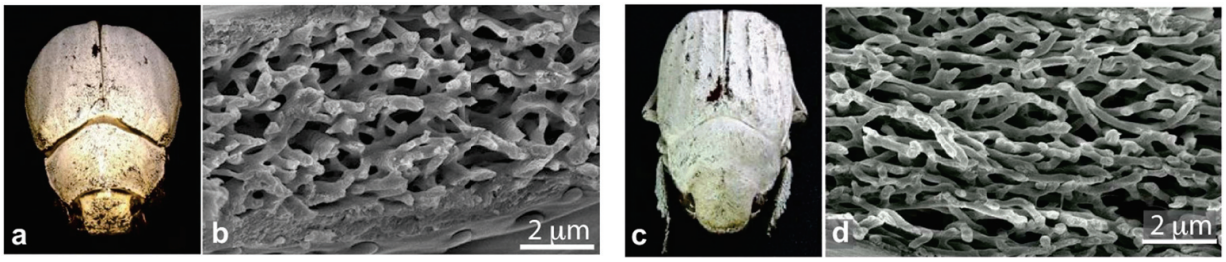


Figure 6.12: Two white beetles and their shell microstructure which has an extraordinary diffuse white reflectance compared to its thickness. Image reprinted with permission (Burresi et al., 2014).

gain insight in the complexity of photonic nanostructures in nature, Karthaus, 2012 is recommended reading. Other books and review articles exist (Parker, 2000; Lafait et al., 2009; Kinoshita, 2008; Fudouzi, 2011; Yu et al., 2013; Wiersma, 2013; Sun et al., 2013) which all aid in understanding these phenomena.

As the epigraph to this chapter suggest, the coherent scattering (not light coherence) caused by periodic structures are hard to hide (Natarov et al., 2014), and ideally these should be avoided altogether for most designs. But designing deterministic structures are at the moment more feasible, and that limitation may not be a severe problem as long as the chosen approach allows enough freedom to produce structural colours. *Le mieux est l'ennemi du bien*[4] as Voltaire says.

[4]Perfect is the enemy of good.

7 Appearance of aluminium

The simple, gentle movements of his hands seemed designed to soothe the walking stick, and when he saw the silver regain its milky sheen, his anger toward the noodles, his wife, and the dentures gradually subsided.

Oh Jung-hee, Tonggyong (The Bronze Mirror)

Models for predicting and analysing optical mechanisms giving aluminium its appearance have been investigated in order to understand the causes – and how to alter them towards a desired behaviour. Whereas the deterministic geometries in the preceding chapters were ideal for deeper understanding of light interaction, focus is now shifted towards non-deterministic systems and phenomenological models.

7.1 Rough surfaces

A substantial part of a surface's appearance is due to its roughness (He et al., 1991). Both macroscopic ("surface finish") and microscopic ("light scattering") features play an important – and sometimes surprising (Chung and Shin, 2013) – role in how a surface is perceived. It is therefore important to be able to describe optical properties of rough surfaces in order to model the visual appearance of aluminium.

Modelling reflection, transmission and absorption due to rough interfaces is a complicated task for several reasons: the random behaviour of the surface might be hard to describe; full wave analysis have to be performed on a large ensemble of surface realisations to capture correct behaviour; and many different effects can take place (Simonsen, 2010), making it hard to use approximate models. For these reasons, rough surface scattering and characterisation is still an active area of research (Maradudin, 2007; Simonsen, 2010).

7.2 Scalar rough surface scattering (P5)

Some of the scattering effects that may take place on a rough surface is enhanced backscattering, polarisation dependent scattering and losses due to surface plasmon polaritons (Simonsen, 2010). These theories requires complex analysis of the rough surface, and will

result in a computational heavy model. This is undesired for model fitting and optimisation purposes, and a simpler theory is used instead. The choice is therefore between scalar methods.

Within computer graphics, a wide range of rough surface scattering models exist (Phong, 1975; Cook and Torrance, 1982; Lafortune et al., 1997; Westin et al., 2004; Walter et al., 2007)[1]. These are often developed with speed and visual appealing behaviour in mind, which makes them infeasible for correlating the roughness BRDF with actual measurements. Even though some models are physically based (He et al., 1991; Stam, 1999; Jensen et al., 2001; Stam, 2001; Walter et al., 2007; Pharr and Humphreys, 2010) since computer graphics users require more and more realism (special effects, animation films), the models are still hard to correlate with experimental parameters.

To ensure a fully physically based method for describing reflection from rough interfaces, the recent Generalized Harvey-Shack (GHS) method for rough surfaces (Harvey et al., 2009; Krywonos et al., 2011) is therefore chosen. It is based on a so-called non-paraxial extension of SDT[2], but instead of Fourier transforming a deterministic function, an expression representing the random height distribution of the surface is transformed. This transformation is then normalised to ensure energy conservation, see P5 and Harvey et al. (1999, 2009) and Krywonos et al. (2011). The result is a theory which is simple to implement and seems to predict scattering from fully reflective surfaces and rough dielectric interfaces reasonably well (Harvey et al., 2010; Schröder et al., 2011; Krywonos et al., 2011; Choi and Harvey, 2013). An efficient implementation of the GHS method tailored for purposes in the thesis is found in P5.

7.3 Reflectance measurements

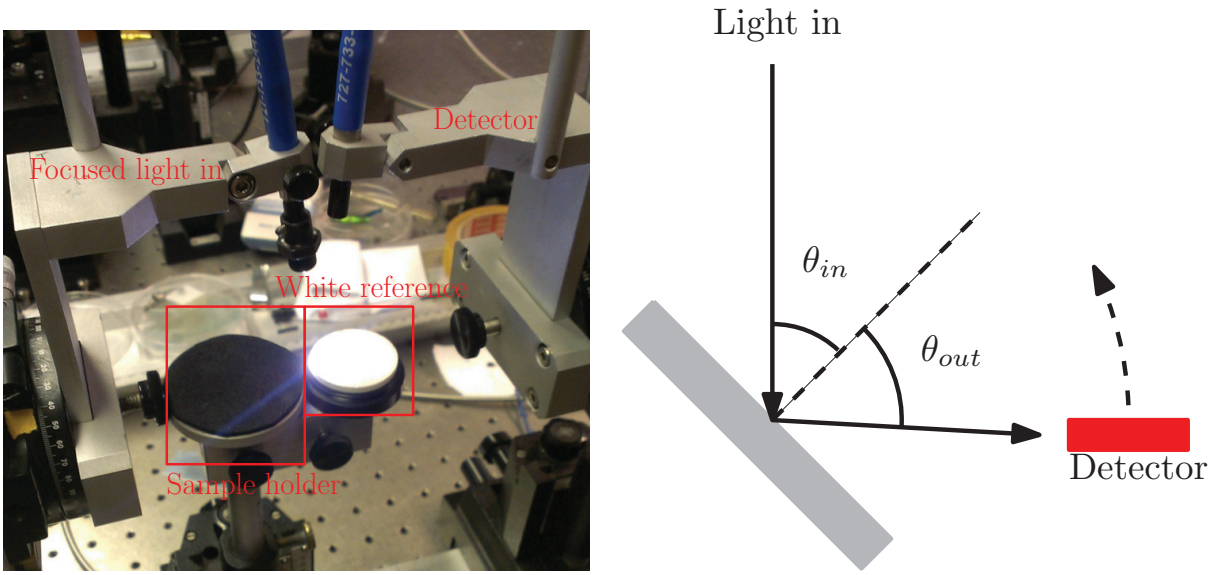
Several rough and polished aluminium surfaces were measured at University of Exeter during the project. One of the purposes was to verify the roughness model in P5. Two different set ups were used for BRDF measurements, and they are briefly explained below. One was based on a white light source and the other on a laser source.

In Figure 7.1(a) a photograph of the white light source set up is seen. A white light source from a 1000 μm fibre is focused on a sample or white reference (for calibration) and the light is collected by a 200 μm fibre connected to a USB2000 Ocean Optics spectrometer. The receiving fibre size gives a high numerical aperture and ensures that the collector extent is small. Both fibres are mounted on rotation stages which makes it possible to adjust incident light angle as well as to scan reflection angles. Both fibres are tilted a bit in order to be able to capture specular reflection at (near-)normal incidence. This set up was also used for the results in P4.

The laser set up is sketched in Figure 7.1(b). Here, a computer controlled motor stage moved a photodetector in a fixed laser and sample set up. The laser light is filtered for

[1]The given references are just few among many and the list is by no mean representative for the whole field.

[2]SDT was introduced in Section 5.1.



(a) (b)

Figure 7.1: BRDF set ups used in the project. (a) a white light based set up. (b) a laser based set up.

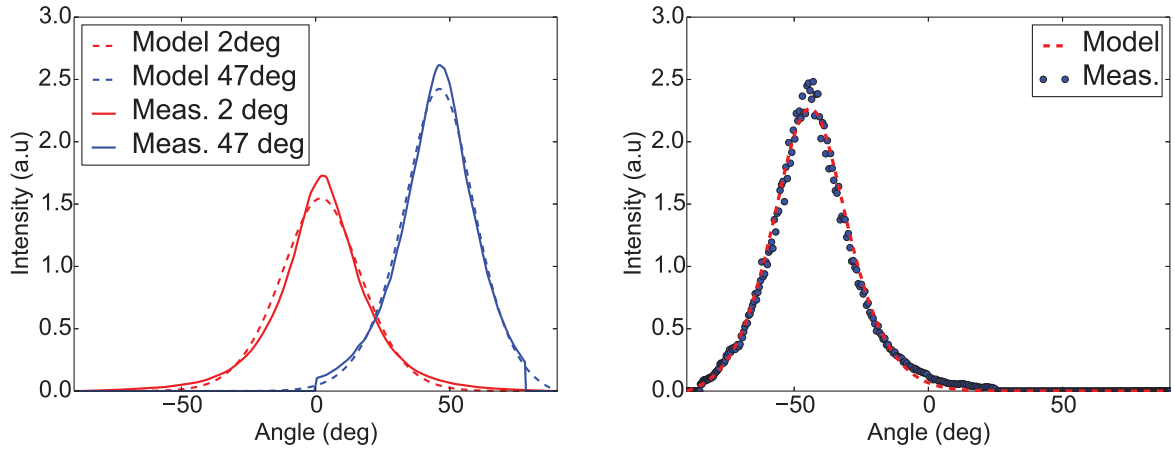
either E or H polarisation. Due to the precision control of the detector and since it is less sensitive to incident light angle than the fibre set up combined with a smaller opening angle, a much finer angular resolution can be achieved in this set up. The drawback is that only one wavelength is covered, and that specular reflection at normal incidence cannot be captured since the detector then would be blocking the light source.

7.4 Rough surface scattering applied to aluminium

Using a Gaussian surface height distribution having a Gaussian autocovariance function (P5), the GHS model was fitted to BRDF measurements for etched aluminium samples. Both the white light source and the laser set up was used. The results for a white light source on an etched, unanodised Al6060 aluminium alloy are seen in Figure 7.2(a). The BRDF was fitted to both incident angles simultaneously, and only using the scattered (and not specular) part of the GHS model. This is because the measured specular reflectance is strongly dependent on the measurement set up. A reasonable fit is seen, which is concluded to be good enough for the purposes of the project.

The same experiment was performed for the laser set up with a result as seen in Figure 7.2(b). Here, an even better fit is seen, which is also expected since the set up has a better resolution. The deviation of the fitted roughness parameters for the two methods is seen to be small.

For modelling aluminium reflectance, it is of course crucial to take the optical behaviour of the aluminium oxide into account as well. Such models are considered in the following sections.



(a) $\sigma = 129.5$ nm, $\ell_c = 1166$ nm (b) $\sigma = 124.3$ nm, $\ell_c = 1083$ nm

Figure 7.2: (a) Reflected intensity for two incident angles of the same surface measured using a focused white light source, and then fitted to the GHS model. (b) H polarised laser measurements and model fit of the same surface with high angular resolution. Besides the rough surface parameters which are described in detail in P5, the intensity scale was also fitted since the measurements were relative.

7.5 The Kubelka-Munk two-stream model

When investigating mechanisms for creating white aluminium, an obvious optical system to consider is that of (white) paint and how it can be modelled. In 1931 Paul Kubelka and Franz Munk proposed a model as "*a first attempt at a theoretical treatment of the optics of coatings*" (Kubelka and Munk, 1931). The method is still widely used today due to its simplicity, but more advanced models based on radiative transfer[3] with less restrictions (Vargas and Niklasson, 1997; Joseph and Thomas, 2012) have to be used to capture more intricate effects[4]. The Kubelka-Munk (KM) equation describes diffuse reflection and transmission from a medium filled with scatters – also called a *turbid* medium – in a plane parallel geometry for incoming diffuse light, see Figure 7.3. Examples of turbid media are paper, paint and milk.

The basic principle of KM is to consider a turbid/scattering medium with a plane interface lit by diffuse light and then quantify how the rate of conversion of diffuse intensity into the forward and backward directions is described for a differential layer in the medium (see again Figure 7.3). For every differential element in the turbid medium, some downwards travelling diffuse radiance I is lost due to scattering S and absorption K. At the same time, some is gained due to backscattering of upwards travelling diffuse radiance J.

[3]Kubelka-Munk theory can be seen as derived from the radiative transfer equation (Thennadil, 2008; Sandoval and Kim, 2014).

[4]In 1971, for example, it was noted that "*On the one hand, the two-flux (Kubelka-Munk) and four-flux calculations are simple enough that they can be done in a few minutes on a desk calculator capable of exponentiation, but their range of reliability is limited*" (Mudgett and Richards, 1971).

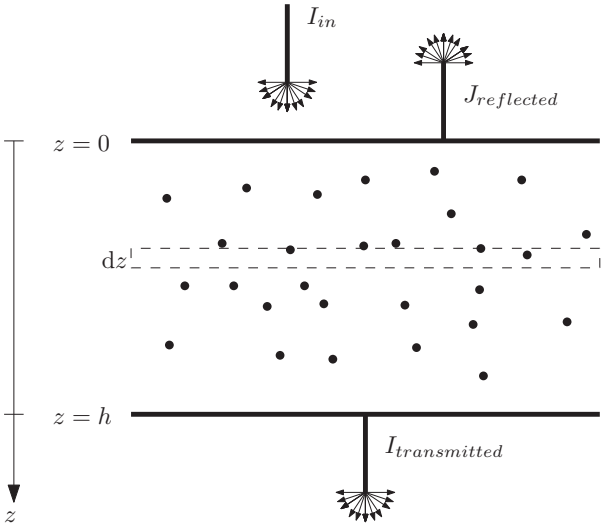


Figure 7.3: Geometry for KM calculations. The basic goal of KM is to calculate how much diffuse intensity is transmitted and reflected for a turbid medium with plane interfaces based on the properties of a differential element $\mathrm{d}z$ of that medium.

Since this process goes both ways, radiances in each differential element in the coating are changed according to the following equations:

$$\frac{\mathrm{d}I}{\mathrm{d}z} = -(S+K)I + SJ, \tag{7.1a}$$

$$\frac{\mathrm{d}J}{\mathrm{d}z} = (S+K)J - SI. \tag{7.1b}$$

This coupled system of first order ordinary differential equations (ODE) can be solved in many ways yielding solutions of superposed exponential functions. A convenient way to write the solution for $S > 0$ is (Murphy, 2006)

$$I(z) = C_1 e^{-Sbz} + C_2 e^{Sbz}, \tag{7.2a}$$

$$J(z) = C_1(a-b)e^{-Sbz} + C_2(a+b)e^{Sbz}, \tag{7.2b}$$

where $a = (S+K)/S$, $b = \sqrt{a^2-1}$ and C_1, C_2 are constants determined by the boundary conditions of the problem.

Albedo

For paint on an absorbing (black) surface, the boundary conditions for the KM equations are

$$I(0) = I_{in}, \tag{7.3a}$$

$$J(h) = 0. \tag{7.3b}$$

Solving (7.2) with those boundary conditions gives

$$C_1 = I_{in}\frac{a+b}{a+b+(b-a)e^{-2Sbh}}, \tag{7.4}$$

$$C_2 = I_{in}\frac{(b-a)e^{-2Sbh}}{a+b+(b-a)e^{-2Sbh}}. \tag{7.5}$$

The ratio of incident to reflected diffuse radiance for a given coating height is then given by

$$R = \frac{J(0)}{I_{in}} = \frac{C_1(a-b)e^0 + C_2(a+b)e^0}{I_{in}} = \frac{(1-e^{-2Sbh})(a^2-b^2)}{a+b+(b-a)e^{-2Sbh}} = \frac{1-e^{-2Sbh}}{a(1-e^{-2Sbh})+b(1+e^{-2Sbh})}. \tag{7.6}$$

This reflection can be referred to as *albedo*. A 100% albedo corresponds to a perfect, diffuse white surface. The albedo for an infinite thick layer is easily derived as

$$R_\infty = \lim_{h\to\infty} R = \frac{1}{a+\sqrt{a^2-1}}. \tag{7.7}$$

Paint albedos

For paint based on TiO_2 particles, the scattering parameter might be around $S = f3\ \mu\text{m}^{-1}$ (Auger et al., 2003), where f is a filling fraction[5] of particles in the paint. The absorption for paint is close to zero, e.g. $K = 10^{-5}\ \mu\text{m}^{-1}$.

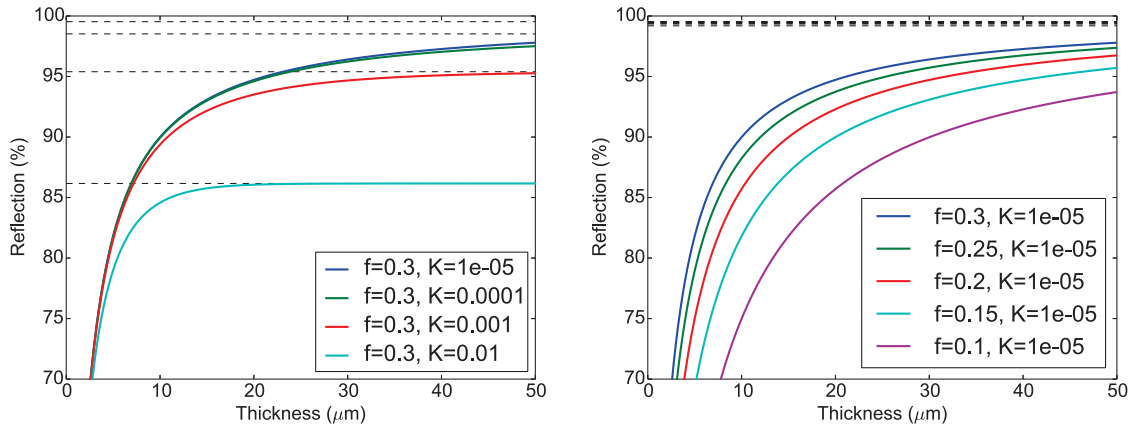
If absorption is present in such a system, for example due to impurities in the paint, even little absorption can greatly affect the overall achievable properties of the paint[6]. This is illustrated in Figure 7.4(a), where albedos for different layer thicknesses of a paint with 30% volume fraction is plotted for different absorption coefficients. Another important property is that the volume fraction of the scattering TiO_2 particles in paint proportionally affects the layer thickness needed to obtain a certain albedo (or coverage). That is, half the volume fraction means double the height to acquire same coverage. Both examples in Figure 7.4 furthermore illustrates how hard it is to obtain the last percentage points of reflection for a paint system. It is therefore important to optimise scattering parameters of a paint system (or aluminium surface) in order to obtain a good albedo in practice.

7.6 The modified Kubelka-Munk model (P6)

The Kubelka-Munk equation does not take interface scattering into account. In order to increase the precision of the model, this was proposed in Murphy (2006), and in P6 this approach is used to estimate reflectance properties of aluminium oxide with incorporated TiO_2 paint particles. The calculations are further backed by a physical simulation where paint particle in a polyurethane matrix are spin coated onto aluminium substrates. Even though particle agglomerations are present in the matrix, trends between experiment and analysis are similar. The model in P6 is a central element in predicting reasonable volume fractions and oxide thicknesses in order to create white aluminium. Later in the project, TiO_2 particles were successfully incorporated in the oxide, and the results are presented in Section 8.2.

[5] Valid for $f \lesssim 30\%$.

[6] That is, R_∞ decreases.



(a) (b)

Figure 7.4: (a) Reflection of paint with $S = f3\ \mu\text{m}^{-1}$ for different values of absorption K. The dashed lines indicate R_∞. (b) The same set up, but with varying f.

7.7 Particle scattering (P6)

Calculating scattering properties of spherical particles can be done using Mie theory (Mie, 1908; Bohren and Huffman, 1983). Since the Kubelka-Munk equation parameters are phenomenological, there is no unambiguous way to convert electromagnetic scattering properties to S and K, but for particles in several size regimes, different conversions are proposed (Brinkworth, 1972; Auger et al., 2003; Thennadil, 2008).

In P6 it is investigated how particle size distributions of different commercially available TiO_2 paint particles will affect the scattering effect if they are incorporated into anodised aluminium. This was done by simulating scattering from single spheres of varying sizes using Mie theory. A rule of thumb is that this single scattering theory is valid for well-dispersed particles up until around 20% or 30% percent, depending on the wanted precision (DuPont, 2007). Using a software called `mstm`[7] (Mackowski, 2014), scattering coefficients of $S \approx f4\ \mu\text{m}^{-1}$ and no absorption $K = 0$ were calculated in most of the visible spectrum. These simple calculations have proven indispensable in identifying viable parameters, when trying to design white aluminium by incorporation of TiO_2 particles.

Multiple particle scattering

There are several cases where considering the scattering of single particles are not sufficient and their electromagnetic interaction needs to be taken into account. During the project, the influence of multiple scattering at a wavelength of 550 nm was calculated for TiO_2 particles with a diameter of 210 nm dispersed in a resin with a refractive index of 1.5[8].

[7]For quick online calculations, see `http://omlc.org/software/mie/`.

[8]Had the simulations been done today, the chosen refractive index would probably be slightly higher, in order to match aluminium oxide better.

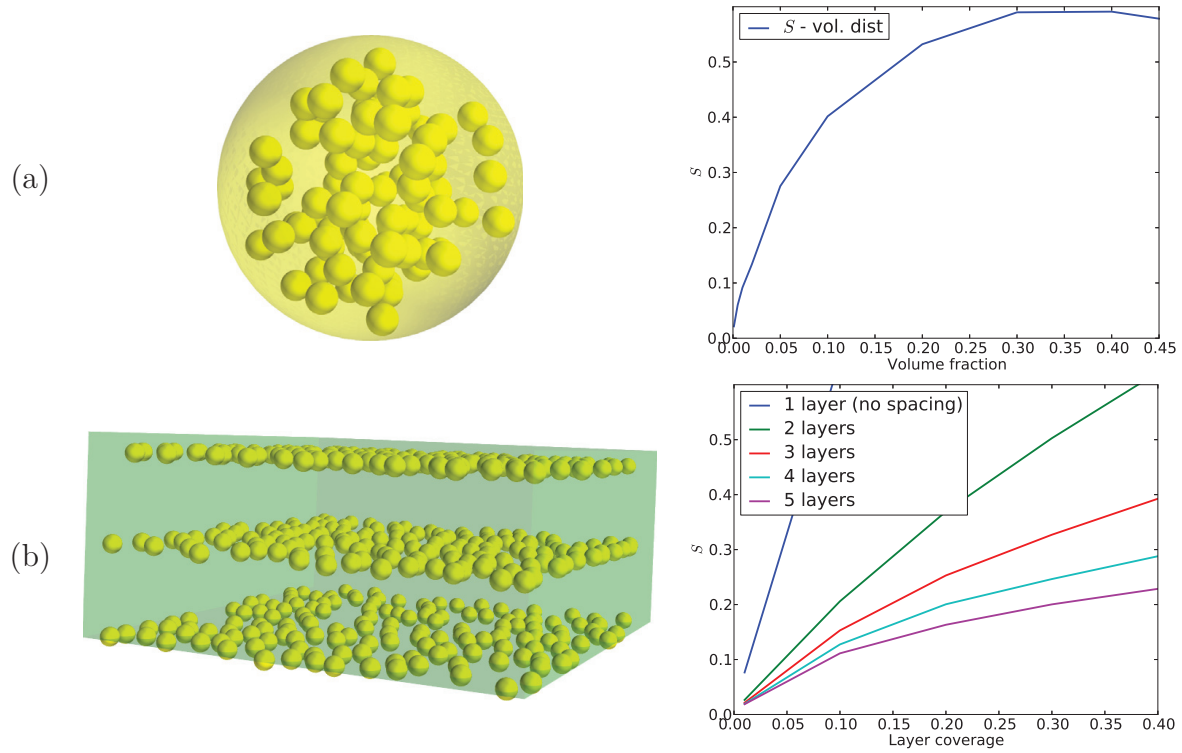


Figure 7.5: (a) Monte Carlo simulations of scattering parameter S@550 nm for a system of 210 nm TiO_2 particles for varying volume fraction. (b) Scattering parameter for the same set up, except that the particles here are ordered in layers with 1 μm spacing. Due to the "volume waste" in between layers, the average scattering efficiency is low.

The wavelength was chosen to be representative for visible light scattering. Several simulations were averaged, and the results are presented in Figure 7.5(a). The found scattering parameter is slightly lower than for the single scattering case (as expected), and it shows that going much above the single scattering regime does not improve scattering much. Furthermore, at these high percentages, the mechanical properties of the oxide are expected to degrade, and ODAAS therefore have no interest in higher values. The calculations were performed using `mstm`, and the method was verified using results from (Auger et al., 2008).

Spray deposition of TiO_2 particles in layers and anodisation in between was proposed during the course of the ODAAS project. It was numerically predicted how many layers would be needed to get a good scattering efficiency. The results are presented in Figure 7.5(b). The graph shows, that for several layers, the overall scattering efficiency per volume is low. Using the KM equations it was predicted that more than 100 layers were needed in order to obtain a strong scattering effect. It was therefore decided not to pursuit that approach to obtain white aluminium.

7.8 Radiative transfer

The KM model is limited to pure diffuse light, and can as such not give angle resolved information on the behaviour of turbid media. This is needed for more detailed analysis. For this purpose, radiative transfer is used (Chandrasekhar, 1960). The radiative transfer equation (RTE) can be written as

$$-\mu\frac{\mathrm{d}u(\tau,\mu,\phi)}{\mathrm{d}\tau} = -u(\tau,\mu,\phi) + \frac{\Omega}{4\pi}\int_{-1}^{1}\int_{0}^{2\pi} p(\mu',\phi',\mu,\phi)u(\tau,\mu',\phi')\mu'\mathrm{d}\mu'\mathrm{d}\phi', \tag{7.8}$$

where u is the (angle resolved) radiance at an optical depth $\tau \propto z$, Ω is the albedo and p is the phase function, which is a function that describes the scattering properties of the turbid medium. The equation can be understood the same way as the governing KM differential equations: for a differential element of the turbid medium, a radiance redistribution takes place due to the scattering properties of the turbid medium.

The RTE has to be discretised in order to be solved. In this project it was done using the Discrete Ordinates Method (DOM) (Chandrasekhar, 1960; Stamnes et al., 1988; Thynell, 1998), which is a widely used approach for solving the RTE – especially for slab geometries[9]. Using DOM and proper descriptions of the phase functions and rough interfaces, it is hopefully possible to describe the optical scattering properties of anodised aluminium. See Figure 7.6 for a graphical interpretation. The model makes detailed comparison across alloys, anodisation condition and pre-/post-treatments possible, thus enabling identification of scattering/absorption properties depending on these choices. This will give knowledge to the ODAAS project on alloy and treatment selection based on optical characteristics. It is expected that the model will work best for oxide where the pores are sealed by hot water since that decreases the anisotropic scattering from the pores which will require an extension of the classical DOM.

Several samples have been prepared to investigate how a RT model can be fitted to anodised aluminium. The current status is described in the following section.

7.9 Optical behaviour of anodised aluminium

Using the laser based measurement set up presented in Section 7.3, several aluminium samples where investigated in order to map and understand differences in their visual appearance. In particular, a description of the scattering properties of the oxide is sought. To investigate the oxide, etched and polished aluminium samples were anodised at 15 V to different oxide thicknesses: 0 μm (no ano.); $\sim$ 3 μm; $\sim$ 10 μm; $\sim$ 17 μm. The samples were subsequently sealed. The purpose of the samples is to fit angle resolved measurements to a RT model and by this extract effective oxide scattering parameters. The measurements were carried out on three different alloys: Al6060 which is a very pure substrate; Al2024 which is less pure; and PA853 which is a commercial alloy used by B&O. Furthermore,

[9] Also referred to as parallel plane geometries.

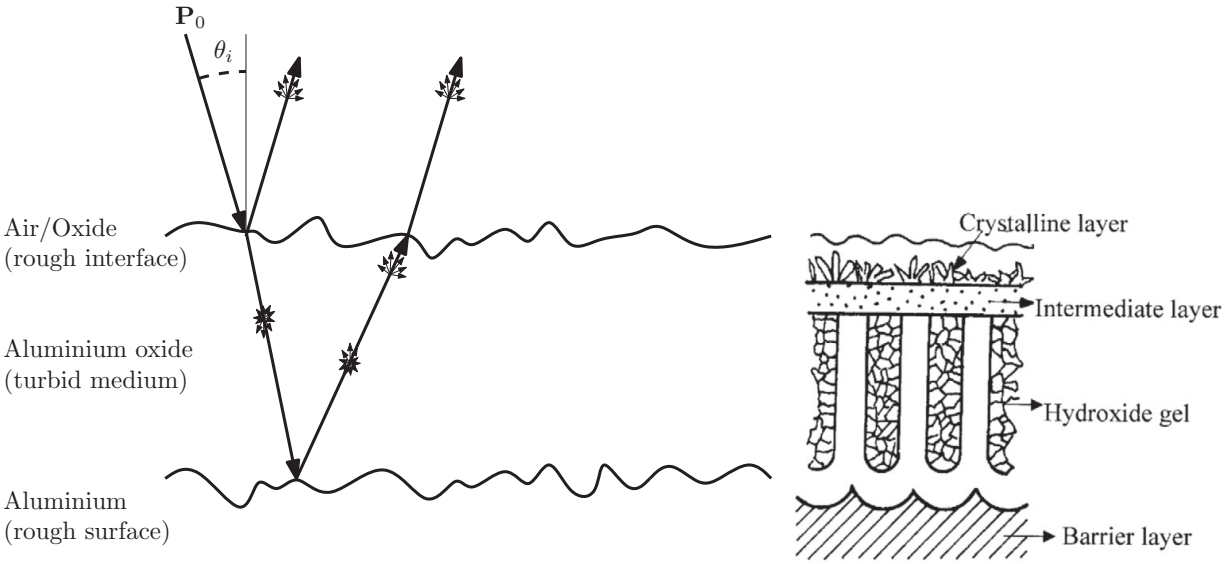


Figure 7.6: Anodised aluminium can optically be described by its rough air/oxide interface, its scattering properties of the oxide and its rough oxide/aluminium interface. Even though several scattering mechanisms are present in the oxide, it is expected that effective scattering parameters relatively independent of oxide height can be extracted. Especially for sealed aluminium, where the pore structures are filled, and the anisotropic scattering is expected to be smaller.

samples were anodised at 10 V and 20 V to a thickness of 10 μm in order to be able to compare differences in anodisation voltage as well.

A selection of the photographed samples is seen in Figure 7.7. Considering the polished samples, it is seen that specular reflection is decreased after anodisation, and that there is an oxide thickness dependence. Also, the sharpness of the mirrored die is strongly dependent on anodisation voltage. Furthermore, the reflection in the pure aluminium (Al6060) seems to be less distorted by the oxide than for the two others. The reason for these changes are expected to be caused by intermetallic particles found in the oxide after anodisation (Aggerbeck et al., 2014a); changes in surface/interface roughness properties, albeit not drastically (Aggerbeck et al., 2014a); and the voltage dependent quality of the pore structure, which even after sealing is expected to give rise to scattering. Considering the etched samples, less variation is observed. The most distinct effect is for the PA853 alloy where specular reflection is better maintained. The reason is probably that the alloy is less rough after being etched under the same conditions as the other alloys.

Unquestionably there are more scattering mechanism present in anodised aluminium

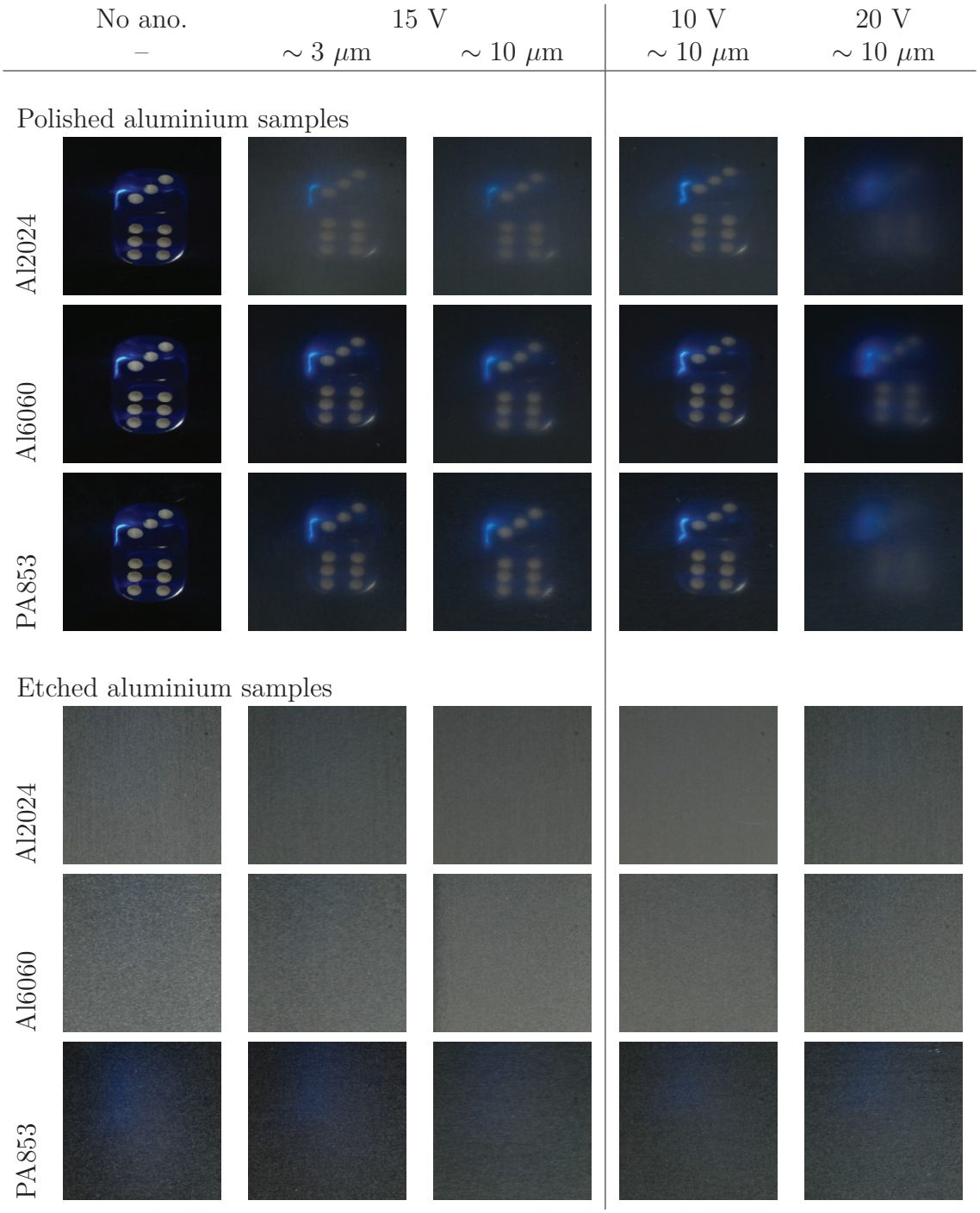


Figure 7.7: Photographs of different three different aluminium alloys taken using the set up in Figure 3.4. The samples have been anodised to different oxide heights using different anodisation voltages.

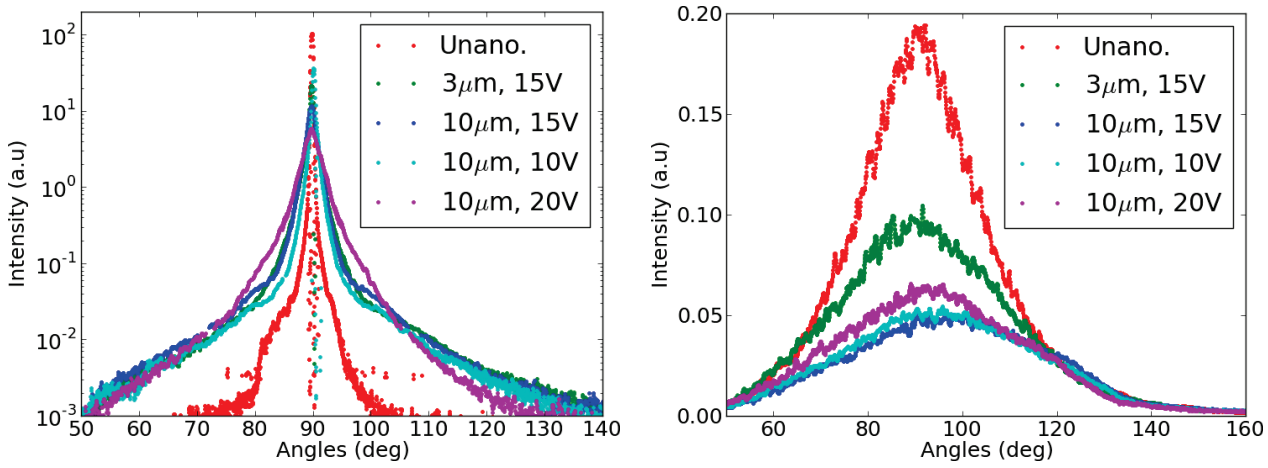


Figure 7.8: Angle resolved reflectance measurements of some of the Al6060 samples in Figure 7.7 using the laser setup presented in Section 7.3. (a) Reflection from specular samples at different heights. Note the logarithmic scale. (b) Reflections from diffuse samples at different heights.

than those just mentioned, but producing a solely physics based model for anodised aluminium seems impossible at the moment. For some special cases (Woltgens et al., 2001; Goossens et al., 2009; Graeve et al., 2011) it has been performed and the work definitely confirm an interest in understanding, predicting and tailoring visual appearance of oxidised aluminium and metals. Much more common is it to find reports in the literature on observation of (changes in) appearance which are not supported by an optical model (Gils et al., 2003; Barchi et al., 2010; Ma et al., 2013; Mirzaei and Bahrololoom, 2014; Aggerbeck et al., 2014b). For these reasons, work was carried out in order to create a framework which should make it easier to predict the appearance of anodised aluminium based on earlier results and RT analysis. Unfortunately the measured angle resolved reflection of the samples (see Figure 7.8) can at the moment of writing be fitted to many different parameters depending on initial guess, and the model is therefore not useful yet.

8 Creating white aluminium

men mørket er hvidt, siger børnene, paradismørket er hvidt, men ikke på den måde hvidt som en kiste er hvid, hvis kister da findes, og ikke på den måde hvidt, som mælken er hvid, hvis mælken da findes

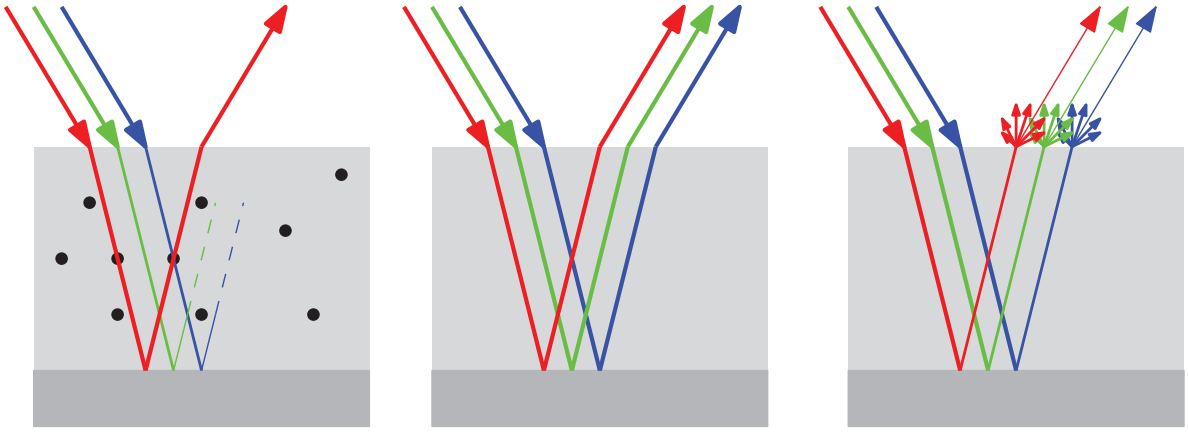
Inger Christensen, Alfabet

The studies presented in the previous chapters have all served as tools or inspiration for proposed methods of creating white aluminium in the ODAAS project. The proposals themselves are not based on elaborate theories since significant gaps still exist between that which can be modelled, that which is desired, and that which can be fabricated. This chapter recapitulates on conceived ideas in the project and comments on their status. It displays how inspiration from the presented investigations has motivated new concepts; it serves as documentation of generated ideas throughout the ODAAS project; and hopefully it will act as inspiration for future work.

The chapter is intentionally kept brief, since the main experimental work has been carried out by other partners in the project. It may therefore be skipped by those with an interest only in the research part of the thesis: it is a demonstration of the outcome of research rather than documentation of the research itself.

8.1 Suggested methods for creating white aluminium

In this section, concepts for white aluminium are presented, and to understand the problems that are addressed, difficulties in realising white aluminium are firstly recapitulated. A white appearance of aluminium can *not* be seen analogous to the coloured appearance of dyed aluminium. It is conceptually very different because both specular and diffuse reflectance are required, as explained in Figure 8.1. Furthermore, for the rendering in Figure 3.3 it was also found that a slight decrease in diffuse reflectivity quickly changes the appearance from clear white to greyish, and it is therefore important to keep the overall reflection high. Even though it has not been investigated in detail, it is most likely also important to preserve surface texture from surface treatment steps such as polishing, extrusion, blasting or similar in order to give a metallic appearance.



(a) Red aluminium (b) Mirror-like aluminium (c) White aluminium concept

Figure 8.1: Illustration of how white aluminium differs from other colours of aluminium. (a) When aluminium is dyed, it gets its colour by absorbing all other part of the visible spectrum, thus making all reflected objects appear – in this case – red. (b) This principle does not work for white, since white is a combination of all colours in the spectrum, and the hue of reflected objects will therefore *not* change. Since this effect is associated with mirrors, a white equivalent to coloured aluminium does *not* exist. (c) From the rendering in Figure 3.3 it is proposed that a combination of diffuse and specular scattering which make darker reflected objects appear whiter than they actually are is the best way to obtain an illusion of white metal.

This means that a successful method should focus on

1. creating diffuse white scattering;
2. partially retaining a specular reflection;
3. preserve surface structure;
4. getting as much light energy reflected back to the observer as possible.

Besides the point of preserving surface texture – which is hard to predict and might be handled by post treatments – all proposed methods address one or several of the issues. In the following sections the methods and a short status on the progress in the project are presented.

Incorporating scattering particles in the oxide

White paint is made by having strongly scattering particles in the binder, giving an overall white diffuse reflection, See Section 7.5. The same scattering mechanism is proposed for anodised aluminium. To scatter effectively, the particles should be well dispersed and have a mean diameter around 0.25 μm and only little variation in particle size (DuPont,

2007). It is unlikely that a strong diffuse scattering and specular mirror effect can co-exist using this approach, unless polishing of the top oxide layer creates enough desired specular reflection. This issue might be covered by combining the approach with some of the later proposals.

Status on modelling Optimal particle sizes, volume concentrations and oxide thicknesses have been predicted from the method presented in P6. The modelling is not precise enough to predict whether or not a mirror effect like the specular samples in Figure 7.7 can be retained. This might be predicted better by a radiative transfer model in the future.

Status on experimental work A lot of work has been done in trying to incorporate TiO_2 particles in the aluminium oxide using different methods (Gudla et al., 2015). It has proven problematic to get a decent, well-dispersed particle concentration without changing the phase of the particles (from rutile to anatase or amorphous). At the time of writing 10 wt.% well-dispersed particles can be incorporated into the aluminium for subsequent anodisation, which have given a decent white appearance. The samples are discussed in Section 8.2.

Caveats Besides experimental issues which seem to have been mostly solved, it is still an open issue whether or not a specular reflection can be recovered. The total reflection of the current samples still lack five to ten percentage point in order to get a white surface comparable with white paint.

Anisotropic scattering

As mentioned in Section 6.3, anisotropic scattering can increase the perceived brightness as seen in Figure 6.9. To enhance the experienced white diffuse scattering above the Lambertian limit, it is therefore proposed to look into methods that can create an optical rough anisotropic aluminium surface having similar properties. Since metals are perceived as extremely bright, a correctly chosen anisotropic white scattering may even appear metallic without a mirror-like specular reflection. Just like butterfly wings and beetle shields can give this metallic impression without a distinct mirror effect.

The anisotropy may be introduced by oriented anisotropic particles in the oxide or by pretexturing the aluminium before anodisation. The pretexturing can be done by pattern imprint (Novelis, 2012), engraving, or anisotropic etching if possible (e.g. by having a flow in the etching bath).

An interesting secondary application is that, by combining the structure with dye or selective multilayer structures, anisotropy will enable colours even more bright than what currently exists.

Status on modelling Ordinary scattering theories (Harvey et al., 1999) predict that anisotropic roughness will give rise to anisotropic scattering, but it has not been relevant

to carry out any detailed simulations. Given more time it would be interesting to generate an anisotropic BRDF and render it in `pbrt` to evaluate the effect on the visual appearance.

Status on experimental work No experimental work has been carried out.

Caveats It is not yet clear if there is a feasible way to obtain a satisfactory anisotropic surface. Furthermore, the effect on appearance is not evaluated in detail.

Optical brightening

A fundamentally different approach to increase diffuse reflectance beyond the Lambertian limit is to use optical brightening agents (OBA) – also known as fluorescent whitening agents (FWA). These are chemical compounds that work by re-emitting blue light from energy absorbed in the ultraviolet spectrum. This enhances the appearance of a white surface, since adding a bit of blue in general will enhance the effect of white.[1]

Since aluminium oxide already contain pores it should be straight-forward to incorporate the OBA's. The idea was originally proposed by Professor Bill Barnes, University of Exeter, when it was stated that 110% reflected energy is needed in order to create a proper white appearance.

Status on modelling No modelling is needed at the current stage.

Status on experiments It has been confirmed that OBA's can be incorporated in the oxide pores and that an enhanced blue reflection is observed. Compared to the data sheet of the OBA, it seems like the absorption edge is shifted when incorporated into the pores. This is at the moment believed to be caused by small pore sizes, and experiments on larger pore sizes will be carried out.

Caveat Besides the shift in absorption edge observed in the experiments, it is still not evaluated if enough re-emission can be obtained from the system in order to get the desired effect. It is also a problem that indoor lighting technologies currently are undergoing major development. This means that there is no guarantee that indoor lighting will contain a substantial amount of UV light in the future.

Dielectric mirrors

In case of lack of specular reflection, it is proposed to apply a semitransparent dielectric multilayer mirror (Leś and Kuroś, 1977) on top of the oxide. The proposed robust multilayer optimisation methods in Section 5.2 can be used for the purpose. The multilayers can be produced using methods like physical vapour deposition (PVD) or by continuing the anodisation using alternating anodising conditions (Yisen et al., 2011).

[1] Another way of "bleaching" clothes is to add a trace of blue dye to it – this process is known as bluing.

Status on modelling As presented in Section 5.2, modelling and optimisation codes are available and have been used.

Status on experimental work A few samples were produced for calibration using PVD. The track was put on hold due to other activities with higher priority.

Caveats Due to the required layer thickness precision, the approach may be hard to apply to commercial products with curved surfaces and large areas. Wear resistance may also be an issue.

Semi-painted surface

A straight-forward way to approach the need for specular and diffuse reflection is to make microholes in a polished aluminium surface, and then fill them with paint. This will in principle retain some specular reflection as well as give a diffuse component.

Status on modelling By visual observation of computer generated images with a prescribed amount of randomly distributed white pixels on a black background it is possible to estimate a desired filling fraction for the first trial. It was decided that around 70% of the surface should be holes for paint in the first trial.

Status on experimental work Several surfaces with etched holes have been produced. Trials were carried out to fill paint in the holes. The aspect ratio of the holes was unfortunately too low and the paint would not stay in. A silicon wafer with deeper holes was also produced, but here the dimensions were too small. It is therefore still unknown how the visual appearance will be for such an approach.

Caveats Apart from the mentioned experimental problems, such a surface will probably be hard to anodize due to the low aluminium coverage.

8.2 Current status

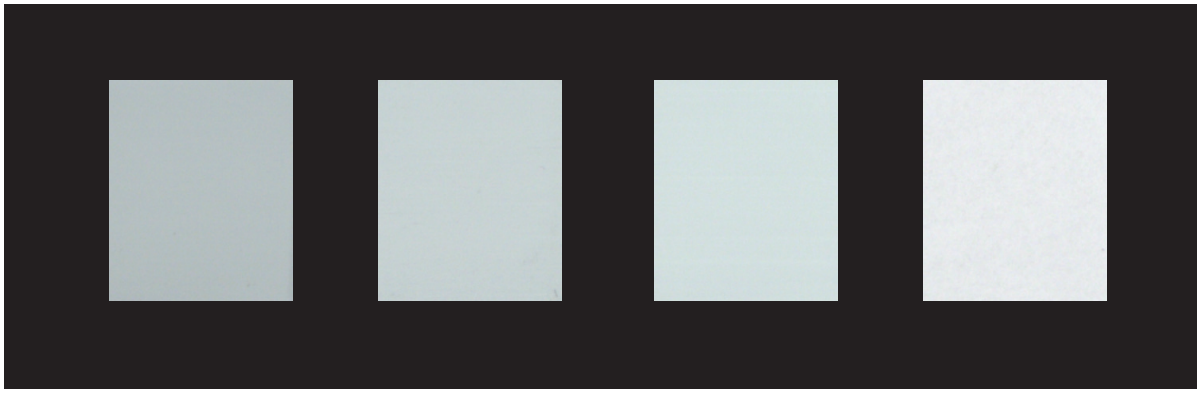
Recent trials on incorporating TiO_2 particles into bulk aluminium and anodising[2] it subsequently have so far made it possible to produce aluminium with an appearance that can be characterised as white. A polished sample anodised to 88 μm can be seen in Figure 8.2[3]. The appearance is diffuse white with a tiny bit of gloss still present. Streaks from polishing are still visible on the surface which helps giving an impression of a metal.

Depending on the surface treatment before anodising and anodising thickness, slightly different effects can be obtained: the anodisation height controls how white the sample

[2]The process steps are unfortunately confidential at the time of writing.

[3]Unfortunately the photo experiment from P7 has not been performed yet.

Figure 8.2: The result of anodising a polished aluminium sample with added TiO_2 particles in the substrate. (image blurred due to confidentiality)



23 μm, blasted 50 μm, grinded 80 μm, polished White paper

Figure 8.3: Comparison of different white aluminium samples with varying anodisation heights and surface treatments. A white paper photographed under similar conditions is shown as a reference.

gets, and surface preparation affects gloss. Cutting out the anodised part of photographies taken under equivalent lighting condition as in Figure 8.2, the different diffuse appearances are compared in Figure 8.3. Due to the camera's white balance the colour representation is precise but not accurate, and a photo of a white piece of paper is therefore presented as a reference. It can be seen that the obtained diffuse white is still lacking in comparison to white paper. It is believed that by tuning the process parameters the appearance can be made whiter than it is now.

Pastel colours

B&O also has an interest in expanding their aluminium palette with pastel colours – as opposed to the bright colouration in Figure 1.1(a). The white aluminium samples have therefore been dyed. Some of the samples are presented in Figure 8.4. The results are promising, and an overview of the different colour trials is presented in Figure 8.5. This additional feature makes the potential of the method even larger.

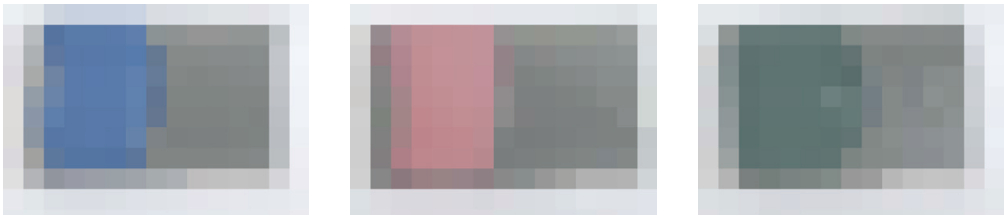


Figure 8.4: Aluminium samples similar to the one in Figure 8.2 where dye has been added before the sealing process. (image blurred due to confidentiality)

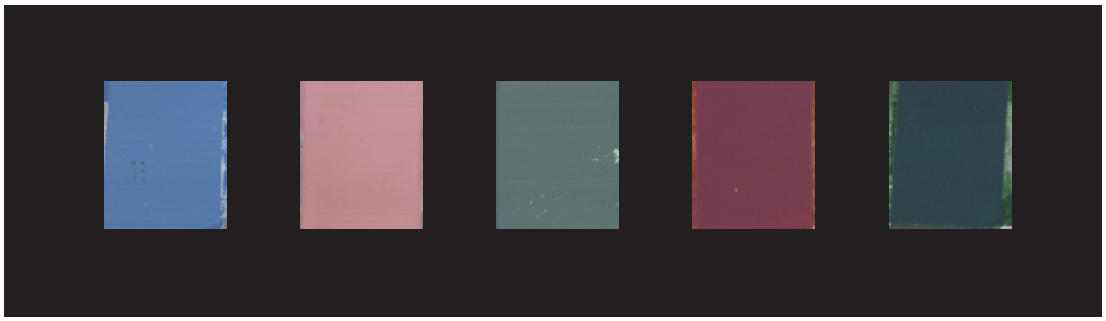


Figure 8.5: Cut outs of different aluminium samples like the ones in Figure 8.4. A novel range of colour appearances can be produced this way.

8.3 Future work

To continue the current experiments such that *an anodised white (milky) appearance* can be obtained, incorporation and anodisation of the particle filled aluminium still needs to be optimised such that the volume fraction of particles is increased to the maximum feasible value. It also has to be investigated if secondary processing steps like incorporation of OBA's or polishing of the anodised layer will enhance the white appearance distinctly. These tasks are of a more practical character, and does not involve new theoretical work.

If those steps cannot give the desired appearance, then other – fundamentally different – concepts have to be tried out, like anisotropic structuring of the surface. At the time of writing it is not believed that this should be necessary in order to satisfy the project goal of ODAAS.

Upscaling laboratory experiments to industrial production and other optimisation steps necessary to have a feasible product are also required. Those tasks are outside the scope of the thesis.

9 Concluding remarks

"Take care," said the count, "you know the saying, my dear sir; a little learning takes you away, a lot brings you back."

Gustave Flaubert, Bouvard and Pécuchet

Interaction between theory and experimental work facilitated the successful fabrication of white aluminium in the ODAAS project. The theoretical understanding obtained from the work in this thesis assisted in determining the visual appearance properties required for a white metallic appearance. Optical models were then used to propose and discard methods for obtaining said appearance. After having selected a particle scattering based solution, critical parameters (particle size and distribution, oxide thickness) were also determined from optical models. Most of the remaining tasks in ODAAS now lie within upscaling and process tuning.

The optical models in the thesis extend beyond anodised aluminium analysis and the appended articles contribute to a general understanding of structural colour design. The optimisation based design approaches in (P1, P2) can lead to realisations of novel colour appearances and design tailored for industrial production (Friis and Sigmund, 2012). The random analysis approach in P3 is a prerequisite for structural colours based on unit structure designs. Furthermore, applications span beyond structural colours and, for example, photovoltaics and transparent display technologies (Hsu et al., 2014) are other research areas that can benefit from the work in this thesis.

Still, the thesis barely scratches the surface of structural colours. The work mainly focused on methods suitable for obtaining white aluminium. Other successful structural colouration methods like plasmon based colour selective surfaces (Clausen et al., 2014; Tan et al., 2014) have therefore not been investigated here. Altogether, structural colour research is growing and future challenges include making structural colours usable for the industry (in order to save paint, create longer lasting colours or obtain new colour appearances).

For this, detailed control of surface appearance, robust designs with viable tolerances and production methods and 3D modelling and virtual prototyping (like computer rendering) are all required. By virtual prototyping, effects like anisotropy and angle dependence can be considered already in the design phase and robust designs ensure a successful transfer from theory to product. The prospects are alluring and this thesis is a good platform to advance further within the field.

Bibliography

Aage, N., N. A. Mortensen, and O. Sigmund (2010). 'Topology optimization of metallic devices for microwave applications'. *International Journal for Numerical Methods in Engineering* **83**, pp. 228–48.

Aggerbeck, M. (2014). 'Tailored Aluminium based Coatings for Optical Appearance and Corrosion Resistance'. PhD thesis. Section for Materials and Surface Engineering, Department of Mechanical Engineering, Technical University of Denmark.

Aggerbeck, M., S. Canulescu, K. Dirscherl, V. E. Johansen, S. Engberg, J. Schou, and R. Ambat (2014a). 'Appearance of anodised aluminium: Effect of alloy composition and prior surface finish'. *Surface & Coatings Technology* **254**, pp. 28–41.

Aggerbeck, M., A. Junker-Holst, D. V. Nielsen, V. C. Gudla, and R. Ambat (2014b). 'Anodisation of sputter deposited aluminium–titanium coatings: Effect of microstructure on optical characteristics'. *Surface & Coatings Technology* **254**, pp. 138–44.

Ahn, J. S., T. R. Hendricks, and I. Lee (2007). 'Control of Specular and Diffuse Reflection of Light Using Particle Self-Assembly at the Polymer and Metal Interface'. *Advanced Functional Materials* **17**, pp. 3619–25.

Ahn, S. H. and L. J. Guo (2009). 'Large-Area Roll-to-Roll and Roll-to-Plate Nanoimprint Lithography: A Step toward High-Throughput Application of Continuous Nanoimprinting'. *ACS Nano* **3**(8), pp. 2304–10.

Auger, J.-C., R. G. Barrera, and B. Stout (2003). 'Scattering efficiency of clusters composed by aggregated spheres'. *Journal of Quantitative Spectroscopy and Radiative Transfer* **79–80**, pp. 521–31.

Auger, J.-C., V. A. Martinez, and B. Stout (2008). 'Theoretical study of the scattering efficiency of rutile titanium dioxide pigments as a function of their spatial dispersion'. *Journal of Coatings Technology and Research* **6**(1), pp. 89–97.

Balanis, C. A. (2012). *Advanced Engineering Electromagnetics*. 2nd. Wiley.

Ball, P. (2005). *Butterflies shine brighter by design*. Nature News.

Bao, G., D. C. Dobson, and K. Ramdani (2002). 'A constraint on the maximum reflectance of rapidly oscillating dielectric gratings'. *SIAM Journal on Control and Optimization* **40**(6), pp. 1858–66.

Barchi, L., U. Bardi, S. Caporali, M. Fantini, A. Scrivani, and A. Scrivani (2010). 'Electroplated bright aluminium coatings for anticorrosion and decorative purposes'. *Progress in Organic Coatings* **67**, pp. 146–51.

Bendsøe, M. P. and N. Kikuchi (1988). 'Generating Optimal Topologies in Structural Design Using a Homogenization Method'. *Computer Methods in Applied Mechanics and Engineering* **71**, pp. 197–224.

Bendsøe, M. P. and O. Sigmund (2004). *Topology Optimization – Theory, Methods and Applications*. Springer.

Billmeyer, F. W. and M. Saltzman (2000). *Billmeyer and Saltzman's principles of color technology*. Ed. by R. S. Berns. 3rd. Wiley.

Blender3D (2014). *blender.org*. `www.blender.org`.

Bohren, C. F. (2001). *Clouds in a Glass of Beer: Simple Experiments in Atmospheric Physics*. 1st. Dover Publications.

Bohren, C. F. and D. R. Huffman (1983). *Absorption and Scattering of Light by Small Particles*. Wiley.

Brinkworth, B. J. (1972). 'Interpretation of the Kubelka-Munk Coefficients in Reflection Theory'. *Applied Optics* **11**(6), pp. 1434–5.

Burresi, M., L. Cortese, L. Pattelli, M. Kolle, P. Vukusic, D. S. Wiersma, U. Steiner, and S. Vignolini (2014). 'Bright-White Beetle Scales Optimise Multiple Scattering of Light'. *Scientific Reports* **4**(6075), pp. 1–7.

Campa, A., J. Krc, and M. Topic (2013). 'Two approached for inchoerent propagation of light in rigorous numerial simulations'. *Progress In Electromagnetics Research* **137**, pp. 187–202.

Çapoğlu, İ. R., C. A. White, J. D. Rogers, H. Subramanian, A. Taflove, and V. Backman (2011). 'Numerical simulation of partially coherent broadband optical imaging using the finite-difference time-domain method'. *Optics Letters* **36**(9), pp. 1596–98.

Chandrasekhar, S. (1960). *Radiative Transfer*. Dover Publications.

Choi, N. and J. E. Harvey (2013). 'Numerical validation of the generalized Harvey–Shack surface scatter theory'. *Optical Engineering* **52**(11), pp. 115103–14.

Christiansen, A. B., J. S. Clausen, N. A. Mortensen, and A. Kristensen (2014). 'Injection moulding antireflective nanostructures'. *Microelectronic Engineering* **121**, pp. 47–50.

Chung, K. and J. H. Shin (2013). 'Range and stability of structural colors generated by *Morpho*-inspired color reflectors'. *Journal of the Optical Society of America A* **30**(5), pp. 962–68.

Chung, K., S. Yu, C.-J. Heo, J. W. Shim, S.-M. Yang, M. G. Han, H.-S. Lee, Y. Jin, S. Y. Lee, N. Park, and J. H. Shin (2012). 'Flexible, Angle-Independent, Structural Color Reflectors Inspired by Morpho Butterfly Wings'. *Advanced Materials (Weinheim, Germany)* **24**, pp. 2375–9.

Clausen, J. S., E. Højlund-Nielsen, A. B. Christiansen, S. Yazdi, M. Grajower, H. Taha, U. Levy, A. Kristensen, and N. A. Mortensen (2014). 'Plasmonic Metasurfaces for Coloration of Plastic Consumer Products'. *Nano Letters* **14**, pp. 4499–504.

Cook, R. L. and K. E. Torrance (1982). 'A Reflectance Model for Computer Graphics'. *ACM Transactions on Graphics* **1**(1), pp. 7–24.

Disney, W. (2014). *Walt Disney Animation Studios – BRDF Explorer*. `www.disneyanimation.com/technology/brdf.html`.

DuPont (2007). *DuPont™ Ti-Pure® titanium dioxide – Titanium dioxide for coatings*.

Dutré, P., K. Bala, and P. Bekaert (2006). *Advanced Global Illumination*. 2nd. A K Peters/CRC Press.

Fairman, H. S., M. H. Brill, and H. Hemmendinger (1997). 'How the CIE 1931 color-matching functions were derived from Wright-Guild data'. *Color Research & Application* **22**(1), pp. 11–23.

Feynman, R. P. (1988). *QED: The Strange Theory of Light and Matter*. 2nd. Princeton Science Library.

Friis, K. S. and O. Sigmund (2012). 'Robust topology design of periodic grating surfaces'. *Journal of the Optical Society of America B* **29**(10), pp. 2935–43.

Fudouzi, H. (2011). 'Tunable structural color in organisms and photonic materials for design of bioinspired materials'. *Science and Technology of Advanced Materials* **12**(064704), p. 7.

Gils, S. V., S. Holten, E. Stijns, M. Vancaldenhoven, H. Terryn, and L. Mattsson (2003). 'Electropolishing of aluminium: processing and assessment of visual appearance'. *Surface and Interface Analysis* **35**, pp. 121–27.

Goossens, V., N. Gotzen, S. V. Gils, E. Stijns, G. V. Assche, R. Finsy, and H. Terryn (2009). 'Predicting reflections of thin coatings'. *Surface & Coatings Technology* **204**, pp. 551–7.

Graeve, I. D., P. Laha, V. Goossens, R. Furneaux, D. Verwimp, E. Stijns, and H. Terryn (2011). 'Colour simulation and prediction of complex nano-structured metal oxide films Test case: Analysis and modeling of electro-coloured anodized aluminium'. *Surface & Coatings Technology* **205**, pp. 4349–54.

Gudla, V. C., F. Jensen, A. Simar, R. Shabadi, and R. Ambat (2015). 'Friction stir processed Al–TiO2 surface composites: Anodising behaviour and optical appearance'. *Applied Surface Science* **324**, pp. 554–62.

Harvey, J. E. (2012). 'Radiometry Rocks! (alternative title: Diffracted Radiance leads to a Generalized Surface Scatter Theory and a Complete Systems Engineering Analysis of Image Quality)'. In: *Tribute to William Wolfe*. Ed. by M. G. Turner. Vol. 8483. Proceedings of SPIE, pp. 848304–9.

Harvey, J. E., N. Choi, and A. Krywonos (2009). 'Calculating BRDFs from Surface PSDs for Moderately Rough Optical Surfaces'. In: *Optical Manufacturing and Testing VIII*. Ed. by J. H. Burge, O. W. Fähnle, and R. Williamson. Vol. 7426. Proceedings of SPIE, pp. 742601–9.

Harvey, J. E., N. Choi, and A. Krywonos (2010). 'Scattering from Moderately Rough Interfaces between Two Arbitrary Media'. In: *Conference on Optical System Contamination – Effects, Measurements, and Control*. Vol. 7794. Proceedings of SPIE, pp. 77940V1–11.

Harvey, J. E. and A. Krywonos (2006). 'Radiance: The Natural Quantity for Describing Diffraction and Propagation'. In: *The Nature of Light: Light in Nature*. Ed. by K. Creath. Proceedings of SPIE, pp. 628503–12.

Harvey, J. E., C. L. Vernold, A. Krywonos, and P. L. Thompson (1999). 'Diffracted radiance: a fundamental quantity in nonparaxial scalar diffraction theory.' *Applied Optics* **38**(31), pp. 6469–81.

Hassan, E., E. Wadbro, and M. Berggren (2014). 'Topology Optimization of Metallic Antennas'. *IEEE Transactions on Antennas and Propagation* **62**(5), pp. 2488–2500.

He, X. D., K. E. Torrance, F. X. Sillion, and D. P. Greenberg (1991). 'A Comprehensive Physical Model for Light Reflection'. In: *SIGGRAPH '91*, pp. 175–86.

Hecht, E. (2001). *Optics*. 4th. Addison-Wesley.

Hooke, R. (1665). *Micrographia*. `www.gutenberg.org`.

Hsu, C. W., B. Zhen, W. Qiu, O. Shapira, B. G. DeLacy, J. D. Joannopoulos, and M. Soljačić (2014). 'Transparent displays enabled by resonant nanoparticle scattering'. *Nature Communications* **5**(3152), pp. 4152–6.

Huygens, C. (1690). *Treatise on Light (org.: Traité de la lumière)*. `gutenberg.org`.

Jensen, H. W., S. R. Marschner, M. Levoy, and P. Hanrahan (2001). 'A Practical Model for Subsurface Light Transport'. In: *SIGGRAPH '01*. ACM SIGGRAPH 2001, pp. 511–18.

Jensen, J. S. and O. Sigmund (2011). 'Topology optimization for nano-photonics'. *Laser & Photonics Reviews* **5**(2), pp. 308–21.

Joseph, R. I. and M. E. Thomas (2012). 'How Accurate Is the Kubelka-Munk Theory of Diffuse Reflection? A Quantitative Answer'. In: *Reflection, Scattering, and Diffraction from Surfaces III*. Ed. by L. M. Hanssen. Vol. 8495, pp. 84950I–9.

Karthaus, O., ed. (2012). *Biomimetics in Photonics*. CRC Press.

Kientz, B., A. Ducret, S. Luke, P. Vukusic, T. Mignot, and E. Rosenfeld (2012). 'Glitter-Like Iridescence within the Bacteroidetes Especially *Cellulophaga spp.*: Optical Properties and Correlation with Gliding Motility'. *Public Library of Science ONE* **7**(12), pp. 1–12.

Kinoshita, S. (2008). *Structural Colors in the Realm of Nature*. World Scientific Publishing Company.

Kinoshita, S. (2013). *Bionanophotonics: An Introductory Textbook*. 1st. Pan Stanford Publishing.

Kinoshita, S., S. Yoshioka, Y. Fujii, and N. Okamoto (2002). 'Photophysics of Structural Color in the *Morpho* Butterflies'. In: *Forma*. Vol. 17, pp. 103–121.

Kolle, M., A. Lethbridge, M. Kreysing, J. J. Baumberg, J. Aizenberg, and P. Vukusic (2013). 'Bio-Inspired Band-Gap Tunable Elastic Optical Multilayer Fibers'. *Advanced Materials (Weinheim, Germany)* **25**, pp. 2239–45.

Kolle, M., P. M. Salgard-Cunha, M. R. J. Scherer, F. Huang, P. Vukusic, S. Mahajan, J. J. Baumberg, and U. Steiner (2010). 'Mimicking the colourful wing scale structure of the *Papilio blumei* butterfly'. *Nature Nanotechnology* **5**, pp. 511–15.

Krywonos, A., J. E. Harvey, and N. Choi (2011). 'Linear systems formulation of scattering theory for rough surfaces with arbitrary incident and scattering angles'. *Journal of the Optical Society of America A* **28**(6), pp. 1121–38.

Kubelka, P. and F. Munk (1931). 'An Article on Optics of Paint Layers (org.: translated Ein Beitrag zur Optik der Farbanstriche, translated by Stephen H. Westin)'. *Zeitschrift für technische Physik* **12**, pp. 593–601.

Lafait, J., S. Berthier, C. Andraud, V. Reillon, and J. Boulenguez (2009). 'Physical colors in cultural heritage: Surface plasmons in glass'. *Comptes Rendus Physique* **10**, pp. 649–59.

Lafortune, E. P. F., S.-C. Foo, K. E. Torrance, and D. P. Greenberg (1997). 'Non-Linear Approximation of Reflectance Functions'. In: *SIGGRAPH '97*, pp. 117–26.

Larouche, S. and L. Martinu (2008). 'OpenFilters: open-source software for the design, optimization, and synthesis of optical filters'. *Applied Optics* **47**(13), pp. C219–30.

Lazarov, B. S., M. Schevenels, and O. Sigmund (2012). 'Topology optimization with geometric uncertainties by perturbation techniques'. *International Journal for Numerical Methods in Engineering* **90**, pp. 1321–36.

Lee, R. T. and G. S. Smith (2009). 'Detailed electromagnetic simulation for the structural color of butterfly wings'. *Applied Optics* **48**(21), pp. 4177–90.

Leś, Z. and J. Kuroś (1977). 'Method for the synthesis of semitransparent wide band dielectric mirrors'. *Thin Solid Films* **46**, pp. 117–26.

Li, Y.-Y. and H.-C. Lee (2014). 'DLAB: a class of daylight-based uniform color space'. *Journal of the Optical Society of America A* **31**(8), pp. 1876–1885.

Licinio, P., M. Lerotic, and M. S. S. Dantas (1999). 'Diffraction by disordered gratings and the Debye–Waller effect'. *American Journal of Physics* **67**(11), pp. 1013–16.

Luukkonen, O., S. I. Maslovski, and S. A. Tretyakov (2011). 'A Stepwise Nicolson–Ross–Weir-Based Material Parameter Extraction Method'. *IEEE Antennas and Wireless Propagation Letters* **10**, pp. 1295–98.

Ma, Y., X. Zhou, G. E. Thompson, J.-O. Nilsson, M. Gustavsson, and A. Crispin (2013). 'Anodizing of AA6063 aluminium alloy profiles: Generation of dark appearance'. *Surface and Interface Analysis* **45**, pp. 1479–84.

Mackowski, D. W. (2014). *scatcodes – MSTM Version 3.0: April 2013*. `www.eng.auburn.edu/users/dmckwski/scatcodes/`.

Maier, S. A. (2010). *Plasmonics: Fundamentals and Applications*. Springer.

Mandel, L. and E. Wolf (1995). *Optical Coherence and Quantum Optics*. 1st. Cambridge University Press.

Maradudin, A. A., ed. (2007). *Light Scattering and Nanoscale Surface Roughness*. Nanostructure Science and Technology. Springer Science.

Matusik, W., H. Pfister, M. Brand, and L. McMillan (2003). 'A data-driven reflectance model'. In: *SIGGRAPH '03*. ACM SIGGRAPH '01, pp. 759–769.

MaxwellRender (2014). www.maxwellrender.com/.

MERL (2014). *Mitsubishi Electric Research Laboratories*. www.merl.com/brdf/.

Mie, G. (1908). 'Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen'. *Annalen der Physik* **330**(3), pp. 377–445.

Mirzaei, M. and M. E. Bahrololoom (2014). 'Influence of pulse currents on the nanostructure and color absorption ability of colored anodized aluminum'. *Vacuum* **99**, pp. 277–83.

Mudgett, P. S. and L. W. Richards (1971). 'Multiple Scattering Calculations for Technology'. *Applied Optics* **10**(7), pp. 1485–502.

Murphy, A. B. (2006). 'Modified Kubelka–Munk model for calculation of the reflectance of coatings with optically-rough surfaces'. *Journal of Physics D: Applied Physics* **39**, pp. 3571–81.

Natarov, D. M., M. Marciniak, R. Sauleau, and A. I. Nosich (2014). 'Seeing the order in a mess: optical signature of periodicity in a cloud of plasmonic nanowires'. *Optics Express* **22**(23), pp. 28190–98.

Nicolson, A. M. and G. F. Ross (1970). 'Measurement of the Intrinsic Properties of Materials by Time-Domain Techniques'. *IEEE Transactions on Instrumentation and Measurement* **19**(4), pp. 377–82.

Novelis (2012). 'The development of a novel coil-to-coil method of engineering the surface topography of rolled products'. In: *VI Aluminium Surface Science & Technology*.

Okada, N., D. Zhu, D. Cai, J. B. Cole, M. Kambe, and S. Kinoshita (2012). 'Rendering *Morpho* butterflies based on high accuracy nano-optical simulation'. *Journal of Optics* **42**(1), pp. 25–36.

OpenFilters (2014). larfis.polymtl.ca/index.php/en/links/openfilters.

O'Shea, D. C., T. J. Suleski, A. D. Kathman, and D. W. Prather (2003). *Diffractive Optics: Design, Fabrication, and Test*. SPIE Publications.

Otomori, M., T. Yamada, J. Andkjær, K. Izui, S. Nishiwaki, and N. Kogiso (2013). 'Level Set-Based Topology Optimization for the Design of an Electromagnetic Cloak With Ferrite Material'. *IEEE Transactions on Magnetics* **49**(5), pp. 2081–4.

Parker, A. R. (2000). '515 million years of structural colour'. *Journal of Optics A: Pure and Applied Optics* **2**, R15–28.

pbrt (2014). *Physically Based Rendering: From Theory to Implementation*. http://en.wikipedia.org/wiki/Color.

Pedrotti[4] (2006). *Introduction to Optics*. 3rd. Addison-Wesley.

Pharr, M. and G. Humphreys (2010). *Physically Based Rendering: From Theory To Implementation*. 2nd. Morgan Kaufmann.

Phong, B. T. (1975). 'Illumination for Computer Generated Pictures'. *Communications of the ACM* **18**(6), pp. 311–17.

Pouya, C., D. G. Stavenga, and P. Vukusic (2011). 'Discovery of ordered and quasi-ordered photonic crystal structures in the scales of the beetle Eupholus magnificus'. *Optics Express* **19**(12), pp. 11355–64.

Röhrig, M., M. Schneider, G. Etienne, F. Oulhadj, F. Pfannes, A. Kolew, M. Worgull, and H. Hölscher (2013). 'Hot pulling and embossing of hierarchical nano- and microstructures'. *Journal of Micromechanics and Microengineering* **23**(105014), pp. 1–7.

Saito, A., Y. Miyamura, Y. Ishikawa, and J. M. andMegumi Akai-Kasaya andYuji Kuwahara (2009). 'Reproduction, Mass-production, and Control of the *Morpho*-butterfly's Blue'. In: *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics II*. Vol. 7205, pp. 720506–9.

Saito, A., M. Yonezawa, J. Murase, S. Juodkazis, V. Mizeikis, M. Akai-Kasaya, and Y. Kuwahara (2011). 'Numerical Analysis on the Optical Role of Nano-Randomness on the *Morpho* Butterfly's Scale'. *Journal of Nanoscience and Nanotechnology* **11**, pp. 2785–92.

Saito, A., S. Yoshioka, and S. Kinoshita (2004). 'Reproduction of the *Morpho* butterfly's blue: arbitration of contradicting factors'. In: *Optical Systems Degradation, Contamination, and Stray Light: Effects, Measurements, and Control*. Ed. by P. T. C. Chen, J. C. Fleming, and M. G. Dittman. Vol. 5526, pp. 188–94.

Sandoval, C. and A. D. Kim (2014). 'Deriving Kubelka–Munk theory from radiative transport'. *Journal of the Optical Society of America A* **31**(3), pp. 628–36.

Schevenels, M., B. S. Lazarov, and O. Sigmund (2011). 'Robust topology optimization accounting for spatially varying manufacturing errors'. *Computer Methods in Applied Mechanics and Engineering* **200**, pp. 3613–27.

Schröder, S., A. Duparré, L. Coriand, A. Tünnermann, D. H. Penalver, and J. E. Harvey (2011). 'Modeling of light scattering in different regimes of surface roughness'. *Optics Express* **19**(10), pp. 9820–35.

Shiomi, S., M. Miyake, and T. Hirato (2012). 'Electrodeposition of Bright Al-Zr Alloy Coatings from Dimethylsulfone-Based Baths'. *Journal of The Electrochemical Society* **159**(4), pp. D225–29.

Simonsen, I. (2010). 'Optics of surface disordered systems'. *The European Physical Journal Special Topics* **181**, pp. 1–103.

Smith, G. S. (2009). 'Structural color of Morpho butterflies'. *American Journal of Physics* **77**(11), pp. 1010–19.

Song, B., S. C. Eom, and J. H. Shin (2014). 'Disorder and broad-angle iridescence from *Morpho*-inspired structures'. *Optics Express* **22**(16), pp. 19386–400.

Stam, J. (1999). 'Diffraction Shaders'. In: *SIGGRAPH' 99*, pp. 101–10.

Stam, J. (2001). 'An illumination model for a skin layer bounded by rough surfaces'. In: *Rendering Techniques 2001*. Ed. by S. J. Gortler and K. Myszkowski, pp. 39–52.

Stamnes, K., S.-C. Tsay, W. Wiscombe, and K. Jayaweera (1988). 'Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media'. *Applied Optics* **27**(12), pp. 2502–09.

Steindorfer, M. A., V. Schmidt, M. Belegratis, B. Stadlober, and J. R. Krenn (2012). 'Detailed simulation of structural color generation inspired by the *Morpho* butterfly'. *Opt. Express* **20**(19), pp. 21485–94.

Stratton, J. A. and L. J. Chu (1939). 'Diffraction Theory of Electromagnetic Waves'. *Physical Review* **56**, pp. 99–107.

Sun, J., B. Bhushan, and J. Tong (2013). 'Structural coloration in nature'. *Royal Society of Chemistry Advances* **3**, pp. 14862–89.

Takai, T., H. Nakao, and F. Iwata (2014). 'Three-dimensional microfabrication using local electrophoresis deposition and a laser trapping technique'. *Optics Express* **22**(23), pp. 28109–17.

Tan, S. J., L. Zhang, D. Zhu, X. M. Goh, Y. M. Wang, K. Kumar, C.-W. Qiu, and J. K. W. Yang (2014). 'Plasmonic Color Palettes for Photorealistic Printing with Aluminum Nanostructures'. *Nano Letters* **14**, pp. 4023–29.

Thennadil, S. N. (2008). 'Relationship between the Kubelka–Munk scattering and radiative transfer coefficients'. *Journal of the Optical Society of America* **25**(7), pp. 1480–5.

Thynell, S. T. (1998). 'Discrete-ordinates method in radiative heat transfer'. *International Journal of Engineering Science* **36**, pp. 1651–75.

Tikhonravov, A. V. and M. K. Trubetskov (2012). 'Modern design tools and a new paradigm in optical coating design'. *Applied Optics* **51**(30), pp. 7319–32.

Tikhonravov, A. V., M. K. Trubetskov, and G. W. DeBell (1996). 'Application of the needle optimization technique to the design of optical coatings'. *Applied Optics* **35**(28), pp. 5493–508.

Vargas, W. E. and G. A. Niklasson (1997). 'Applicability conditions of the Kubelka–Munk theory'. *Applied Optics* **36**(22), pp. 5580–6.

Vukusic, P., R. Kelly, and I. Hooper (2009). 'A biological sub-micron thickness optical broadband reflector characterized using both light and microwaves'. *Journal of the Royal Society Interface* **6**, pp. 193–201.

Vukusic, P., B. Hallam, and J. Noyes (2007). 'Brilliant Whiteness in Ultrathin Beetle Scales'. *Science* **315**, p. 348.

Vukusic, P. and J. R. Sambles (2003). 'Photonic structures in biology'. *Nature* **424**, pp. 852–55.

Walter, B., S. R. Marschner, H. Li, and K. E. Torrance (2007). 'Microfacet Models for Refraction through Rough Surfaces'. In: *Eurographics Symposium on Rendering*. Ed. by J. Kautz and S. Pattanaik. 195–206.

Wang, F., B. S. Lazarov, and O. Sigmund (2011). 'On projection methods, convergence and robust formulations in topology optimization'. *Structural and Multidisciplinary Optimization* **43**, pp. 767–84.

Wang, W., W. Zhang, J. Gu, Q. Liu, T. Deng, D. Zhang, and H.-Q. Lin (2013). 'Design of a structure with low incident and viewing angle dependence inspired by *Morpho* butterflies'. *Scientific Reports* **3**(3426), p. 9.

Weir, W. B. (1974). 'Automatic Measurement of Complex Dielectric Constant and Permeability at Microwave Frequencies'. *Proceedings of the IEEE* **62**(1), pp. 33–6.

Westin, S. H., H. Li, and K. E. Torrance (2004). 'A Comparison of Four BRDF Models'. In: *Proceedings of the Eurographics Symposium on Rendering*, pp. 1–10.

Wiersma, D. S. (2013). 'Disordered photonics'. *Nature Photonics* **7**, pp. 188–96.

Wiscombe, W. J. (1977). 'The Delta-M Method: Rapid Yet Accurate Radiative Flux Calculations for Strongly Asymmetric Phase Functions'. *Journal of the Atmospheric Sciences* **34**, pp. 1408–22.

Woltgens, H.-W., I. Friedrich, W. K. Njoroge, and W. T. ans Matthias Wuttig (2001). 'Optical, electrical and structural properties of AlTi and AlCr thin films'. *Thin Solid Films* **388**, pp. 237–44.

Worgull, M., M. Schneider, M. Röhrig, T. Meier, M. Heilig, A. Kolew, K. Feit, H. Hölscher, and J. Leuthold (2013). 'Hot embossing and thermoforming of biodegradable three-dimensional wood structures'. *Royal Society of Chemistry Advances* **3**, pp. 20060–64.

Wright, W. D. (1928–29). 'A re-determination of the trichromatic coefficients of the spectral colours'. *Transactions of the Optical Society* **XXX**(4), pp. 141–164.

Yeh, P. (2005). *Optical Waves in Layered Media*. Wiley.

Yisen, L., C. Yi, L. Zhiyuan, H. Xing, and L. Yi (2011). 'Structural coloring of aluminum'. *Electrochemistry Communications* **13**, pp. 1336–39.

Yoshioka, S. (2013). 'Structural Color in Nature – Basic Observations and Analysis'. In: *Pattern Formations & Oscillatory Phenomena*. Ed. by S. Kinoshita. Elsevier.

Yu, K., T. Fan, S. Lou, and D. Zhang (2013). 'Biomimetic optical materials: Integration of nature's design for manipulation of light'. *Progress in Materials Science* **58**, pp. 825–873.

AA proposed SDT based approach to incorporate coherence effects

This appendix shows how spatial coherence in a scalar diffraction theory (SDT) setting can be related to the coherent case by convolution. It is expected that the results to some extend relate to structures which cannot be described by SDT as well. The validity of the approach has only been verified by the author. The proposed method is carried out in 2D for simplicity.

Setup

Having a surface profile with a phase disturbance described by

$$p(x) = A(x)e^{4\pi h(x)/\lambda}, \tag{A.1}$$

its reflected irradiance is proportional to

$$I(x') \propto |(\mathcal{F}p)(x')|^2, \tag{A.2}$$

where $\mathcal{F}$ is the Fourier transform defined in P1. If p is described by a stochastic ergodic process in time, then irradiance can be defined as the time-averaged irradiance over infinite time. Due to the ergodicity of the process, this corresponds to considering an ensemble average of all realizations of p:

$$I(x') := \mathbb{E}(I(x')) = \mathbb{E}\left(|(\mathcal{F}p)(x')|^2\right) = \mathbb{E}\left((\mathcal{F}p)(x')\overline{(\mathcal{F}p)(x')}\right). \tag{A.3}$$

According to Mandel and Wolf (1995, pp. 56–58), the average can be calculated as

$$I(x') = (\mathcal{F}R_{pp})(x'), \tag{A.4}$$

where $R_{pp}(\tau) = \mathbb{E}(\,\overline{p(x)}p(x+\tau)\,)$ is the so-called autocorrelation function of p. This can also be written as

$$R_{pp}(\tau) = (p * \overline{p^b})(\tau), \tag{A.5}$$

where $*$ is convolution and $p^b(x) = p(-x)$ [1].

[1] Note that from this it holds that

$$(\mathcal{F}R_{pp})(x') = \left[\mathcal{F}(p * \overline{p^b})\right](x') = (\mathcal{F}p)(x') \cdot \overline{(\mathcal{F}p)(x')} = |(\mathcal{F}p)(x')|^2, \tag{A.6}$$

Coherence

Consider the modelling of the incoherent of light as a time varying pertubation of the phase function corresponding with respect to position, the phase function is modified as

$$p(x,t) = z(x,t)p_0(x), \tag{A.7}$$

where z is a stochastic process and p_0 is the phase function for the deterministic structure. The time-averaged intensity of the response of p described by (A.4) now requires R_{pp} to be found as

$$\begin{aligned} R_{pp}(\tau) &= \mathbb{E}\left(\overline{p_0(x)z(x)}z(x+\tau)p_0(x+\tau)\right) \\ &= \mathbb{E}\left(\overline{p_0(x)}p_0(x+\tau)\cdot\overline{z(x)}z(x+\tau)\right) \\ &= \mathbb{E}\left(\overline{p_0(x)}p_0(x+\tau)\right)\cdot\mathbb{E}\left(\overline{z(x)}z(x+\tau)\right) \\ &\quad + \operatorname{cov}\left(\overline{p_0(x)}p_0(x+\tau), \overline{z(x)}z(x+\tau)\right) \\ &= R_{p_0p_0}R_{zz} + \operatorname{cov}\left(\overline{p_0(x)}p_0(x+\tau), \overline{z(x)}z(x+\tau)\right). \end{aligned} \tag{A.8}$$

Since covariance between a stochastic and a deterministic function is zero, the above equation simplifies to

$$R_{pp}(\tau) = R_{p_0p_0}R_{zz}, \tag{A.9}$$

from which the irradiance can be calculated as

$$\begin{aligned} I(x') &= (FR_{pp})(x') \\ &= (\mathcal{F}R_{p_0p_0})(x') * (\mathcal{F}R_{zz})(x') \\ &= (\mathcal{F}R_{zz})(x') * |(\mathcal{F}p_0)(x')|^2. \end{aligned} \tag{A.10}$$

This shows that the modelled disturbance of coherence can be taken into account by a convolution.

Numerical verification

An array, p_0 of 160 points have been defined and furthermore a function $z(x) = e^{2\pi i z_0(x)}$ where $z_0(x)$ is a stochastic process which is a sequence of numbers drawn from a uniform distribution and with a specified "correlation radius", R (in points).

The following sizes have then been calculated

$$\begin{aligned} \text{Coherent:} &\quad |(\mathcal{F}p_0)(x')|^2, \\ \text{Real:} &\quad < |(\mathcal{F}zp_0)(x')|^2 > \\ \text{Approx.:} &\quad (\mathcal{F}R_{zz})(x') * |(\mathcal{F}p_0)(x')|^2 \end{aligned} \tag{A.11}$$

which shows the validity of the expression for deterministic functions.

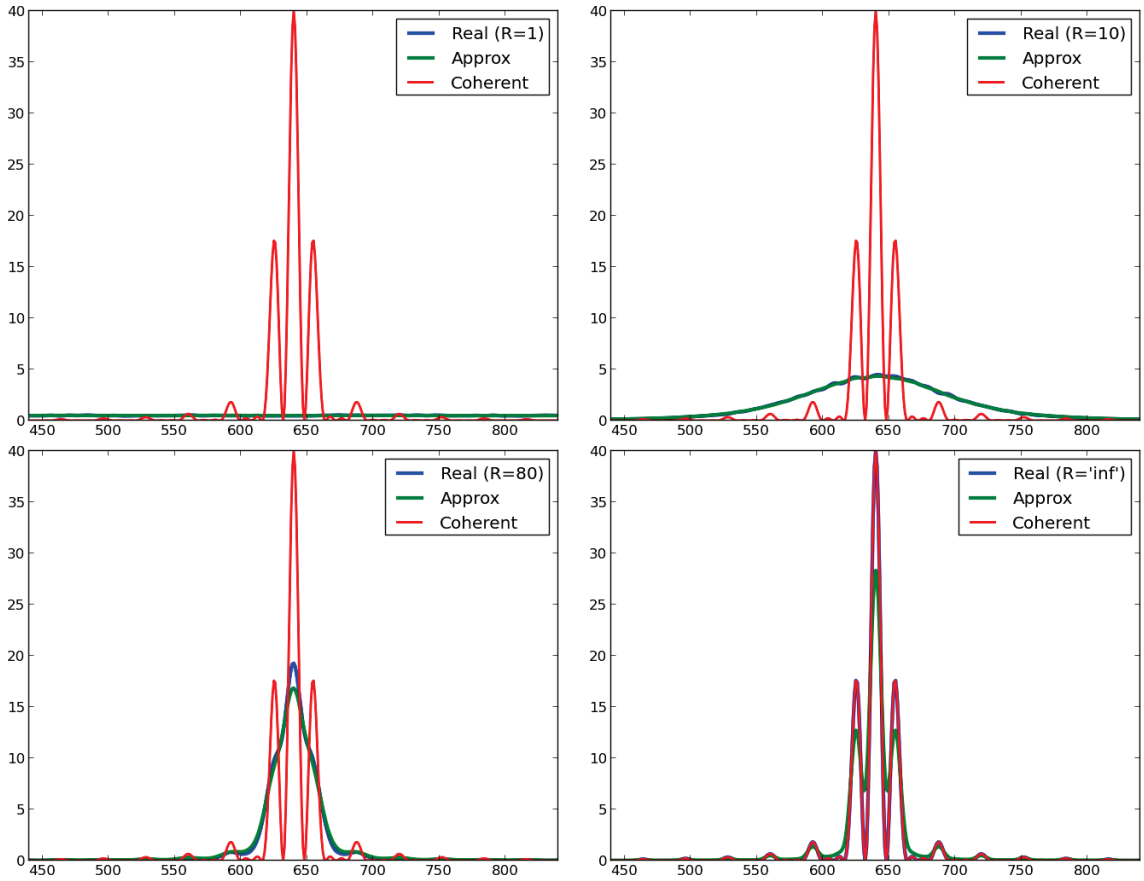


Figure A.1: Numerical coherence experiment based on SDT.

for different R and a selection of results are presented in Figure A.1. The red curve marks the coherent response for a deterministic structure. In case of low coherence (low R), the signature of the coherent response is non-existing. By increasing R, more features from the coherent response emerges, and for $R = \infty$, the coherent response is obtained. It seems like the approximation does not capture this well for large R. This might be because more averages and better numerical resolution are required.

Hopefully this work can be validated and carried on in a context which will make it useful for approximating the effect of coherence for more complex structures.

BNumerical prediction of anodised aluminium reflection

Here follows a derivation of the radiative transfer equation (RTE) implementation used in the Ph.D.for simulating anodised aluminium reflectance. The geometry consists of a slab of scattering/turbid medium with two rough surfaces as boundary conditions. This set up is depicted in Figure B.1.

The numerical approach is based on the so-called Discrete Ordinates Method (DOM) from the 60's (Chandrasekhar, 1960) and its improvement in the 80's (Stamnes et al., 1988). The method is still a widely-used, accurate way of solving radiative transfer problems. The actual implementation is strongly inspired by Stam (2001) which also considers optical properties of rough surfaces. In the roughness model in the article they do not consider propagation of a collimated component, which is assumed when using the Harvey-Shack methods (P5). In this way, the derivation differs from Stam (2001).

The radiative transfer equation

For the slab geometry, the RTE is described by

$$-\mu \frac{\mathrm{d}u(\tau,\mu,\phi)}{\mathrm{d}\tau} = -u(\tau,\mu,\phi) + \mathcal{S}\{u\}(\tau,\mu,\phi), \tag{B.1}$$

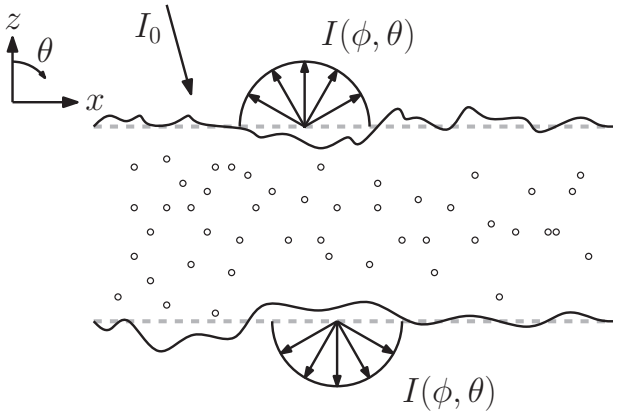


Figure B.1: Geometry for which we want to solve the radiative transfer equation: a scattering/turbid medium surrounded by two rough interfaces.

where u is radiance, $\tau \propto z$ is optical depth and $\pm\omega = (\pm\mu = \pm\cos\theta, \phi)$ is a direction where positive μ indicates an upwards stream and negative μ indicates a downwards stream, $\mathcal{S}$ is a linear operator on u describing the scattering properties of the turbid medium. This could for example be caused by particle scattering.

For this system solution of the following kind is assumed:

$$u = u^c + u^s, \tag{B.2}$$

where u_c is a collimated component (represented by a Dirac delta function) and u^s is a scattered component (represented by finite-valued/bounded function). Substituting the two equations into each other yields

$$-\mu\frac{\mathrm{d}u^c}{\mathrm{d}\tau} - \mu\frac{\mathrm{d}u^s}{\mathrm{d}\tau} = -u^c - u^s + \mathcal{S}\{u^c\} + \mathcal{S}\{u^s\}. \tag{B.3}$$

Since a bounded function and a Dirac delta function are linear independent, the equation can be splitted into two:

$$\mu\frac{\mathrm{d}u^c}{\mathrm{d}\tau} = u^c, \tag{B.4a}$$

$$-\mu\frac{\mathrm{d}u^s}{\mathrm{d}\tau} = -u^s + \mathcal{S}\{u^c\} + \mathcal{S}\{u^s\}. \tag{B.4b}$$

An obvious way to solve this system is then to first to solve (B.4a) and use that solution to solve (B.4b). Which will be done in the following sections.

In all examples, incident light will be assumed collimated and is thus defined by

$$u_{in}(\omega) = I_0\delta(\omega - \omega_i), \tag{B.5}$$

where ω_i is the incoming direction and δ is the Dirac delta function. To simplify the problem further, $\phi = 0$ is assumed [1] meaning that

$$u_{in}(\omega) = I_0\delta(\mu + \mu_i)\delta(\phi). \tag{B.6}$$

where $-\mu_i$ is the incoming direction.

Notation

Before continuing, the following notation and information is introduced to ease derivations.

Direction A superscripted + means that a function is only considered for $\mu > 0$ and a $-$ indicates $\mu < 0$, that is

$$u^{c+} = u^c|_{\mu>0}\,, \tag{B.7a}$$

$$u^{c-} = u^c|_{\mu<0}\,. \tag{B.7b}$$

[1] Without loss of generality as long as isotropic scattering is considered.

Norm A $|\cdot|$ is used to indicate that the scaling of a δ-function is taken, that is

$$|c\delta(\mu-\mu_i)\delta(\phi)| = c. \tag{B.8}$$

Snell's law Consider an incoming collimated light source, say u_{in}. Snell's law state that

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1}, \tag{B.9}$$

which rewritten using μ can be expressed as

$$\begin{aligned}\frac{\sqrt{1-\mu_1^2}}{\sqrt{1-\mu_2^2}} = \frac{n_2}{n_1} \Leftrightarrow \\ \mu_2 = \pm\sqrt{1-(n_1/n_2)^2(1-\mu_1^2)}.\end{aligned} \tag{B.10}$$

Refracted angles in media will then be denoted by a $\cdot'$, meaning that

$$\mu' = \text{sign}(\mu)\sqrt{1-(n_1/n_2)^2(1-\mu^2)}. \tag{B.11}$$

Energy consideration of refracted quantities By varying the direction of the collimated light, energy conservation needs to be considered. Energy, E, of u can be calculated as

$$E = \int_0^{2\pi}\int_{-1}^{1} u(\phi,\mu)\mu \mathrm{d}\phi\mathrm{d}\mu, \tag{B.12}$$

which means that for a collimated beam (say u_{in}),

$$E = \int_0^{2\pi}\int_{-1}^{1} \delta(\mu-\mu_i)\delta(\phi)\mu \mathrm{d}\phi\mathrm{d}\mu = -\mu_i. \tag{B.13}$$

In a refracted medium, obviously $E' = -\mu_i'$ if it is not corrected, and it should therefore be corrected such that

$$|u_{in}'| = \frac{\mu_i}{\mu_i'}|u_{in}|. \tag{B.14}$$

Boxed equations All boxed equations are equations that needs to be implemented in the numerical solution.

Boundary conditions

In order to solve the system, boundary conditions are needed. The following two sections considers the boundary equations for collimated and diffuse light separately.

Boundary condition for collimated light

For collimated light, define $t_{12}^c, r_{21}^c, r_{23}^c$ as the collimated transmittance from medium 1 to 2, the collimated reflectance from medium 2 to 1 and the collimated reflectance from medium 2 to 3. See Figure . If τ_b is the optical depth of the second rough interface, then the boundary conditions can be written as

$$u^{c-}(0) = t_{12}^c \frac{\mu_i}{\mu_i'} |u_{in}| \delta(\mu + \mu_i')\delta(\phi) + r_{21}^c \delta(\mu + \mu_i')\delta(\phi) |u^{c+}(0)|, \tag{B.15a}$$

$$u^{c+}(\tau_b) = r_{23}^c |u^{c-}(\tau_b)| \delta(\mu - \mu_i')\delta(\phi). \tag{B.15b}$$

Boundary condition for scattered light

Using the same approach as for the collimated light, the following boundary conditions can be found:

$$u^{s-}(0) = \mathcal{R}_{21}\{u^{s+}(0)\} + \mathcal{R}_{21}\{u^{c+}(0)\} + \mathcal{T}_{12}\{u_{in}\}, \tag{B.16a}$$

$$u^{s+}(\tau_b) = \mathcal{R}_{23}\{u^{s-}(\tau_b)\} + \mathcal{R}_{23}\{u^{c-}(\tau_b)\}. \tag{B.16b}$$

Here the main differences are that the reflectance/transmittances are now linear operators, since they act on the whole radiated field, and that also the part of the collimated reflectance which is converted into diffuse reflectance needs to be added.

Solution for the collimated light

After having defined the system completely, the collimated RTE in (B.4a) is now solved first. This system is a first order homogeneous ordinary differential equation (ODE) and its solution therefore takes the form

$$u^c = C_0(\mu) e^{\tau/\mu}. \tag{B.17}$$

From the boundary conditions, it is clear that the constant (with respect to τ) will take the following form[2]:

$$C_0(\mu) = \left[C_0^+ \delta(\mu + \mu_i') + C_0^- \delta(\mu - \mu_i') \right] \delta(\phi). \tag{B.18}$$

From this, the two unknown constants can be found using (B.15):

$$C_0^- \delta(\mu + \mu_i')\delta(\phi) = t_{12}^c \frac{\mu_i}{\mu_i'} |u_{in}| \delta(\mu + \mu_i')\delta(\phi) + r_{21}^c \delta(\mu + \mu_i')\delta(\phi) |u^{c+}(0)| \Leftrightarrow$$

$$C_0^- = t_{12}^c \frac{\mu_i}{\mu_i'} |u_{in}| + r_{21}^c C_0^+, \tag{B.19}$$

$$C_0^+ \delta(\mu - \mu_i')\delta(\phi) e^{-\tau/\mu_i'} = r_{23}^c |u^{c-}(\tau_b)| \delta(\mu - \mu_i')\delta(\phi) \Leftrightarrow$$

$$\boxed{C_0^+ = r_{23}^c C_0^- e^{2\tau_b/\mu_i'}.} \tag{B.20}$$

[2]Remember that μ_i' is a negative quantity, since the incoming light is travelling downwards

From these two equations C_0^- can be expressed as

$$C_0^- = t_{12}^c \frac{\mu_i}{\mu_i'}|u_{in}| + r_{21}^c r_{23}^c C_0^- e^{2\tau_b/\mu_i'} \Leftrightarrow$$

$$\boxed{C_0^- = \frac{\mu_i}{\mu_i'} \frac{t_{12}^c}{1 - r_{21}^c r_{23}^c e^{2\tau_b/\mu_i'}}|u_{in}|.} \tag{B.21}$$

The complete solution to (B.17) can now be found by calculating the two boxed equation and substituting the solution back into the system.

Solution for the scattered light

This system cannot be solved analytically. Instead, the equations will be stated in a convenient form for discretization and then solved numerically. Since the diffuse part of the RTE in (B.4b) is inhomogeneous, a particular and a homogeneous solution need to be found.

Rewriting the particular solution

Now that the collimated solution is obtained, the inhomogeneous term in (B.4b) can be rewritten as

$$\begin{aligned}
\mathcal{S}\{u^c\} &= \mathcal{S}\left\{\left[C_0^+\delta(\mu+\mu_i') + C_0^-\delta(\mu-\mu_i')\right]\delta(\phi)e^{\tau/\mu}\right\} \\
&= C_0^+ e^{-\tau/\mu_i'}\mathcal{S}\left\{\delta(\mu+\mu_i')\delta(\phi)\right\} + C_0^- e^{\tau/\mu_i'}\mathcal{S}\left\{\delta(\mu-\mu_i')\delta(\phi)\right\} \\
&\equiv C_0^+ e^{-\tau/\mu_i'} S^{c+} + C_0^- e^{\tau/\mu_i'} S^{c-}.
\end{aligned} \tag{B.22}$$

A particular solution to (B.4b) must then take the form

$$\boxed{u_{par}^s = C_1(\mu)e^{-\tau/\mu_i'} + C_2(\mu)e^{\tau/\mu_i'}.} \tag{B.23}$$

The unknown functions can be found by inserting the solution into the original equation (B.4b):

$$\begin{aligned}
\frac{\mu}{\mu_i'}C_1(\mu)e^{-\tau/\mu_i'} - \frac{\mu}{\mu_i'}C_2(\mu)e^{\tau/\mu_i'} &= -C_1(\mu)e^{-\tau/\mu_i'} - C_2(\mu)e^{\tau/\mu_i'} + C_0^+ e^{-\tau/\mu_i'}S^{c+} \\
&\quad + C_0^- e^{\tau/\mu_i'}S^{c-} + \mathcal{S}\left\{C_1(\mu)e^{-\tau/\mu_i'} + C_2(\mu)e^{\tau/\mu_i'}\right\} \Leftrightarrow \\
&= -C_1(\mu)e^{-\tau/\mu_i'} - C_2(\mu)e^{\tau/\mu_i'} + C_0^+ e^{-\tau/\mu_i'}S^{c+} \\
&\quad + C_0^- e^{\tau/\mu_i'}S^{c-} + \mathcal{S}\left\{C_1(\mu)\right\}e^{-\tau/\mu_i'} + \mathcal{S}\left\{C_2(\mu)\right\}e^{\tau/\mu_i'}.
\end{aligned}$$

In this system the $e^{\tau/\mu_i'}$ and $e^{-\tau/\mu_i'}$ terms are linearly independent, and the equation can therefore be splitted in two:

$$\frac{\mu}{\mu_i'}C_1(\mu) = -C_1(\mu) + C_0^+ S^{c+} + \mathcal{S}\left\{C_1(\mu)\right\} \tag{B.24a}$$

$$-\frac{\mu}{\mu_i'}C_2(\mu) = -C_2(\mu) + C_0^- S^{c-} + \mathcal{S}\left\{C_2(\mu)\right\} \tag{B.24b}$$

Solving these system numerically requires evaluation of the $\mathcal{S}$-operator on the constants. This will be done by discretizing $\mathcal{S}$, which will be performed later.

Rewriting the homogeneous solution

The homogeneous solution does also require evaluation of the scattering and reflectance/transmittance operators, and nothing can be done at the current point to rewrite the equation.

Discrete solution

By expressing the solution as the cosine series (which is possible due to isotropy assumption)

$$u^s(\mu,\phi) = \sum_{k=0}^{\infty} u_k^s(\tau,\mu)\cos k(\phi-\phi_0) \tag{B.25}$$

and truncating it and using the assumption of $\phi_0 = 0$ u^s it can be approximated as

$$u^s(\tau,\mu,\phi) = \sum_{k=0}^{N_f} u_k^s(\tau,\mu)\cos k\phi. \tag{B.26}$$

If p and r_{ij}, t_{ij} are expressed as cosine series as well, then it can be noted that all terms in the Fourier series are linear independent since

$$\int_0^{2\pi} cosk\phi\cos\ell\phi \mathrm{d}\phi = \begin{cases} 2\pi & \text{for } k=\ell=0, \\ \pi & \text{for } k=\ell\neq 0, \\ 0 & \text{for } k\neq\ell, \end{cases} \qquad k,\ell\in\mathbb{Z}. \tag{B.27}$$

That is, all operators can be "linearised".

More notation

Discretisation of μ Since there is a need to integrate over μ, the variable is discretised in its Gauss points to facilitate this (Stam, 2001). It is furthermore discretised in its positive and negative values, since a discontinuity might exists at the boundary, which would give unprecise Gauss integration. So each half space is discretised into M points:

$$\boldsymbol{\mu} = (\mu_1,\mu_2,\ldots,\mu_M,-\mu_1,-\mu_2,\ldots,-\mu_M)^T. \tag{B.28}$$

k subscript Since there now is N_f linearly independent systems to solve, the corresponding field, operator, ..., is denoted by a subscripted k.

Boldface vectors All discretised fields, like u_k, will in the discrete version be represented by a vector denoted $\boldsymbol{u}_k$ such that

$$\boldsymbol{u}_k(\tau) = (u_k(\tau,\mu_1),\ldots u_k(\tau,-\mu_M))^T. \tag{B.29}$$

Boldface operators Discretized operators will also be denoted by boldface letters. The operators are matrices and not vectors.

c superscript - light scattered by collimated light Scattering of collimated light will be indicated by a superscripted c. This mean that e.g. $\boldsymbol{S}^c$ is a vector representing the scattered distribution for some incoming collimated light. In particular, the following discretizations are defined:

$$\boxed{\boldsymbol{T}^c_{in,k} = I_0 \frac{1}{\pi} \int_{-\pi}^{\pi} \mathcal{T}_{12}\{\delta(\boldsymbol{\omega} - \omega_i)\delta(\phi)\} \cos(k\phi) \mathrm{d}\phi,} \tag{B.30}$$

$$\boxed{\boldsymbol{R}^{c+}_{21,k} = C_0^+ e^{-\tau/\mu_i'} \frac{1}{\pi} \int_{-\pi}^{\pi} \mathcal{R}_{21}\{\delta(\boldsymbol{\mu} - \mu_i')\delta(\phi)\} \cos(k\phi) \mathrm{d}\phi,} \tag{B.31}$$

$$\boxed{\boldsymbol{R}^{c-}_{23,k} = C_0^- e^{\tau/\mu_i'} \frac{1}{\pi} \int_{-\pi}^{\pi} \mathcal{R}_{23}\{\delta(\boldsymbol{\mu} - \mu_i')\delta(\phi)\} \cos(k\phi) \mathrm{d}\phi.} \tag{B.32}$$

Determining $\boldsymbol{C}_1, \boldsymbol{C}_2$

Using the introduced notation, the discrete versions of C_1 and C_2 can now be found. Rewriting (B.24) on discrete form for the k'th linearisation they are given by

$$\frac{1}{\mu_i'} \boldsymbol{W} \boldsymbol{C}_{1,k} + \boldsymbol{C}_{1,k} - \boldsymbol{S}_k \boldsymbol{C}_{1,k} = C_0^+ \boldsymbol{S}^{c+} \Leftrightarrow$$

$$\boxed{(\boldsymbol{W}/\mu_i' + \boldsymbol{I} - \boldsymbol{S}_k)\, \boldsymbol{C}_{1,k} = C_0^+ \boldsymbol{S}^{c+}} \tag{B.33a}$$

$$-\frac{1}{\mu_i'} \boldsymbol{W} \boldsymbol{C}_{2,k} + \boldsymbol{C}_{2,k} - \boldsymbol{S}_k \boldsymbol{C}_{2,k} = C_0^- \boldsymbol{S}^{c-} \Leftrightarrow$$

$$\boxed{(-\boldsymbol{W}/\mu_i' + \boldsymbol{I} - \boldsymbol{S}_k)\, \boldsymbol{C}_{2,k} = C_0^- \boldsymbol{S}^{c-}} \tag{B.33b}$$

where

$$\boldsymbol{W} = \mathrm{diag}(\mu_1, \mu_2, \ldots) \tag{B.34}$$

and $\boldsymbol{I}$ is the identity matrix. These linear systems can be solved using standard linear solvers to find $\boldsymbol{C}_1, \boldsymbol{C}_2$.

Determining the homogeneous solution

The discrete solution $\boldsymbol{u}_k$ to the radiative transfer equation can be found by solving

$$-\boldsymbol{W} \frac{\mathrm{d}\boldsymbol{u}_k(\tau)}{\mathrm{d}\tau} = -\boldsymbol{u}_k(\tau) + \boldsymbol{S}_k \boldsymbol{u}_k(\tau) \Leftrightarrow \tag{B.35}$$

$$\frac{\boldsymbol{u}_k(\tau)}{\mathrm{d}\tau} = \boldsymbol{M}_k \boldsymbol{u}_k(\tau), \tag{B.36}$$

where $\boldsymbol{M}_k = \boldsymbol{W}^{-1}(\boldsymbol{I} - \boldsymbol{S}_k)$ and $\boldsymbol{I}$ is the identity matrix. This equation is a first order ordinary matrix differential equation, and its solution can be found by solving an eigenvalue problem and applying boundary conditions. An elegant solution to the eigenvalue problem can be found (Stamnes et al., 1988), and a straight-forward implementation is presented in Stam (2001). For brevity, those calculations are therefore left out, and the results is given as (using the same variables as in Stam (2001))

$$\begin{pmatrix} \boldsymbol{u}^{s+}_{hom,k}(\tau) \\ \boldsymbol{u}^{s-}_{hom,k}(\tau) \end{pmatrix} = \begin{pmatrix} \boldsymbol{V}^+_k & \boldsymbol{V}^-_k \\ \boldsymbol{V}^-_k & \boldsymbol{V}^+_k \end{pmatrix} \begin{pmatrix} \boldsymbol{E}_k(\tau)\boldsymbol{C}^+_{3,k} \\ \boldsymbol{E}_k(-\tau)\boldsymbol{C}^-_{3,k} \end{pmatrix}, \tag{B.37}$$

where $\boldsymbol{E}_k(\tau) = e^{\boldsymbol{L}^+_k \tau}$. And $\boldsymbol{V}^+, \boldsymbol{V}^-$ and $\boldsymbol{L}$ are taken from Stam, 2001. The matrix $\boldsymbol{L}$ is a diagonal matrix.

Boundary conditions

The discrete version of the boundary conditions in (B.16) are

$$\boldsymbol{u}^{s-}_k(0) = \boldsymbol{R}_{21,k}\boldsymbol{u}^{s+}_k(0) + \boldsymbol{R}^{c+}_{21,k}(0) + \boldsymbol{T}^c_{in,k}, \tag{B.38a}$$

$$\boldsymbol{u}^{s+}_k(\tau_b) = \boldsymbol{R}_{23,k}\boldsymbol{u}^{s-}_k(\tau_b) + \boldsymbol{R}^{c-}_{23,k}(\tau_b). \tag{B.38b}$$

where

$$\boldsymbol{u}^s_k = \boldsymbol{u}^s_{hom,k} + \boldsymbol{u}^s_{par,k}. \tag{B.39}$$

By insertion in (B.38a), this can be reformulated as

$$\begin{aligned} \boldsymbol{V}^-_k\boldsymbol{C}^+_{3,k} + \boldsymbol{V}^+_k\boldsymbol{C}^-_{3,k} + \boldsymbol{C}^-_{1,k} + \boldsymbol{C}^-_{2,k} &= \boldsymbol{R}_{21,k}\left(\boldsymbol{V}^+_k\boldsymbol{C}^+_{3,k} + \boldsymbol{V}^-_k\boldsymbol{C}^-_{3,k} + \boldsymbol{C}^+_{1,k} + \boldsymbol{C}^+_{2,k}\right) \\ &\quad + \boldsymbol{R}^{c+}_{21,k}(0) + \boldsymbol{T}^c_{in,k} \Leftrightarrow \\ (\boldsymbol{V}^-_k - \boldsymbol{R}_{21,k}\boldsymbol{V}^+_k)\boldsymbol{C}^+_{3,k} + (\boldsymbol{V}^+_k - \boldsymbol{R}_{21,k}\boldsymbol{V}^-_k)\boldsymbol{C}^-_{3,k} &= \boldsymbol{R}_{21,k}\boldsymbol{C}^+_{1,k} - \boldsymbol{C}^-_{1,k} + \boldsymbol{R}_{21,k}\boldsymbol{C}^+_{2,k} - \boldsymbol{C}^-_{2,k} \\ &\quad + \boldsymbol{R}^{c+}_{21,k}(0) + \boldsymbol{T}^c_{in,k}. \end{aligned} \tag{B.40}$$

By insertion in (B.38b), this can be reformulated as

$$\begin{aligned} &\boldsymbol{V}^+_k\boldsymbol{E}_k(\tau_b)\boldsymbol{C}^+_{3,k} + \boldsymbol{V}^-_k\boldsymbol{E}_k(-\tau_b)\boldsymbol{C}^-_{3,k} + \boldsymbol{C}^+_{1,k}e^{-\tau_b/\mu'_i} + \boldsymbol{C}^+_{2,k}e^{\tau_b/\mu'_i} \\ &\quad = \boldsymbol{R}_{23,k}\left(\boldsymbol{V}^-_k\boldsymbol{E}_k(\tau_b)\boldsymbol{C}^+_{3,k} + \boldsymbol{V}^+_k\boldsymbol{E}_k(-\tau_b)\boldsymbol{C}^-_{3,k} + \boldsymbol{C}^-_{1,k}e^{-\tau_b/\mu'_i} + \boldsymbol{C}^-_{2,k}e^{\tau/\mu'_i}\right) + \boldsymbol{R}^{c-}_{23,k}(\tau_b) \Leftrightarrow \\ &(\boldsymbol{V}^+_k - \boldsymbol{R}_{23,k}\boldsymbol{V}^-_k)\boldsymbol{E}_k(\tau_b)\boldsymbol{C}^+_{3,k} + (\boldsymbol{V}^-_k - \boldsymbol{R}_{23,k}\boldsymbol{V}^+_k)\boldsymbol{E}_k(-\tau_b)\boldsymbol{C}^-_{3,k} \\ &\quad = \boldsymbol{R}_{23,k}\boldsymbol{C}^-_{1,k}e^{-\tau_b/\mu'_i} - \boldsymbol{C}^+_{1,k}e^{-\tau_b/\mu'_i} + \boldsymbol{R}_{23,k}\boldsymbol{C}^-_{2,k}e^{\tau/\mu'_i} - \boldsymbol{C}^+_{2,k}e^{\tau_b/\mu'_i} + \boldsymbol{R}^{c-}_{23,k}(\tau_b). \end{aligned}$$

Defining[3]

$$\boldsymbol{RHS}^+ = \boldsymbol{R}_{21,k}\boldsymbol{C}^+_{1,k} - \boldsymbol{C}^-_{1,k} + \boldsymbol{R}_{21,k}\boldsymbol{C}^+_{2,k} - \boldsymbol{C}^-_{2,k} + \boldsymbol{R}^{c+}_{21,k}(0) + \boldsymbol{T}^c_{in,k}, \tag{B.41}$$

$$\boldsymbol{RHS}^- = (\boldsymbol{R}_{23,k}\boldsymbol{C}^-_{1,k} - \boldsymbol{C}^+_{1,k})e^{-\tau_b/\mu'_i} + (\boldsymbol{R}_{23,k}\boldsymbol{C}^-_{2,k} - \boldsymbol{C}^+_{2,k})e^{\tau_b/\mu'_i} + \boldsymbol{R}^{c-}_{23,k}(\tau_b), \tag{B.42}$$

$$\tilde{\boldsymbol{C}}^+_{3,k} = \boldsymbol{E}_k(\tau_b)\boldsymbol{C}^+_{3,k}, \tag{B.43}$$

[3]The last definition is to scale the system such that it is easier to solve numerically

$\boldsymbol{C}_{3,k}$ can then be found by solving the following system:

$$\boxed{\begin{pmatrix} (\boldsymbol{V}_k^- - \boldsymbol{R}_{21,k}\boldsymbol{V}_k^+)\boldsymbol{E}_k(-\tau_b) & \boldsymbol{V}_k^+ - \boldsymbol{R}_{21,k}\boldsymbol{V}_k^- \\ \boldsymbol{V}_k^+ - \boldsymbol{R}_{23,k}\boldsymbol{V}_k^- & (\boldsymbol{V}_k^- - \boldsymbol{R}_{23,k}\boldsymbol{V}_k^+)\boldsymbol{E}_k(-\tau_b) \end{pmatrix} \begin{pmatrix} \tilde{\boldsymbol{C}}_{3,k}^+ \\ \boldsymbol{C}_{3,k}^- \end{pmatrix} = \begin{pmatrix} \boldsymbol{RHS}^+ \\ \boldsymbol{RHS}^- \end{pmatrix}.} \tag{B.44}$$

The scattered solution can now be written as

$$\boldsymbol{u}_r^s = \boldsymbol{u}^{s+}(0) = \sum_k^N \boldsymbol{V}_k^+ \boldsymbol{E}(-\tau_b)\tilde{\boldsymbol{C}}_{3,k}^+ + \boldsymbol{V}_k^- \boldsymbol{C}_{3,k}^- + \boldsymbol{C}_{1,k}^+ + \boldsymbol{C}_{2,k}^+, \tag{B.45}$$

$$\boldsymbol{u}_t^s = \boldsymbol{u}^{s-}(\tau_b) = \sum_k^N \boldsymbol{V}_k^- \tilde{\boldsymbol{C}}_{3,k}^+ + \boldsymbol{V}_k^+ \boldsymbol{E}_k(-\tau_b)\boldsymbol{C}_{3,k}^- + \boldsymbol{C}_{1,k}^- e^{-\tau/\mu_i'} + \boldsymbol{C}_{2,k}^- e^{\tau/\mu_i'}, \tag{B.46}$$

and the forward and backwards scattered radiance at any place in the turbid medium (or at the interfaces) is now found.

A note on the scattering function

By referring to Stam (2001) for the solution to the homogeneous system, details on how to calculate $\boldsymbol{S}$ were not presented. In Stam (2001), the so-called Henyey-Greenstein phase function was used to represent the scattering behaviour of the turbid medium. This is a very used function, since it can represent strongly forward scattering media as well as diffuse scattering media. This choice might not be the most appropriate for describing scattering in aluminium oxide, but it was chosen as a first approach. If the phase function is kept, the Legendre polynomials used to calculate the coefficient should be done using *fully normalized associated Legendre* functions in order to keep the numerics stable. Whether it is kept or a more appropriate function is found, a δ-DOM (Wiscombe, 1977) implementation should be used to ensure better numerical stability for strongly forward scattering functions.

Status on the numerical implementation

The implementation of the described system without the rough surfaces against the results in, to which they seemed to match. No reference examples for the rough surfaces used in P5 exists, but it was checked that the energy was conserved for the lossless case, and that the single surface reflection cases matched results that could be obtained from P5. The model has been parameter fitted to anodised aluminium reflectance data, but no consistent results have been obtained so far. No gradients were calculated for the parameter fitting – most likely this will improve the accuracy of the method[4].

[4] And it will definitely improve the speed!

Publication P1

Design of structurally colored surfaces based on scalar diffraction theory

Design of structurally colored surfaces based on scalar diffraction theory

Villads Egede Johansen,* Jacob Andkjær, and Ole Sigmund

Technical University of Denmark, Department of Mechanical Engineering, 2800 Kgs. Lyngby, Denmark
**Corresponding author: vejo@mek.dtu.dk*

Received September 12, 2013; accepted October 29, 2013;
posted November 25, 2013 (Doc. ID 197308); published January 6, 2014

In this paper we investigate the possibility of controlling the color and appearance of surfaces simply by modifying the height profile of the surface on a nanoscale level. The applications for such methods are numerous: new design possibilities for high-end products, color engraving on any highly reflective surface, paint-free text and coloration, UV-resistant coloring, etc. In this initial study, the main focus is on finding a systematic way to obtain these results. For now the simulation and optimization is based on a simple scalar diffraction theory model. From the results, several design issues are identified: some colors are harder to optimize for than others, and some can be produced by only a few height levels, whereas others require more complex structures. It is shown that a wide range of results can be obtained.

OCIS codes: (230.1950) Diffraction gratings; (050.2770) Gratings; (050.6624) Subwavelength structures; (240.6700) Surfaces; (330.1690) Color; (330.7326) Visual optics, modeling.
http://dx.doi.org/10.1364/JOSAB.31.000207

1. INTRODUCTION

Structural colors are colors caused by the interaction of light with small structures—most of them containing features comparable to the wavelength of light [1–3]. This makes it possible to obtain new colors (or appearances) that cannot be obtained through absorption-based methods like dyeing. Typical examples of structural colors include CDs, which contain gratings in the micrometer range, thin-film interference in soap bubbles due to their micrometer thickness, and the colorful appearance of peacock feathers due to even more complex microstructural effects [2,4].

Already in 1665 the effect of structural colors was described in the literature by Robert Hooke: "In this we are able from a colourless body to produce several coloured bodies, affording all the variety of Colours imaginable" [5]. This was only a few years after he formulated the wave theory of light, and since then many light experiments and color phenomena have been described. This is particularly true regarding the era after the introduction of Maxwell's equations. Furthermore, the increase in microscope resolution and computational power in recent decades has made it possible to study micro- and nanostructures and their interaction with light on an even more detailed level. This is evident in the detailed study of the microstructure of the Morpho butterfly (see [6–8]).

A lot of effort has been put into these investigations, and it has even been possible to successfully reproduce the color effect of the Morpho butterfly [9]. Few, though, have tried to approach structural colors from a different aspect than biomimicry. In this paper we will consider the inverse problem: if a color effect is desired, which grating structure can represent it? Based only on variations of the surface profile, this paper studies how well color effects can be obtained using gradient-based optimization of structures analyzed using scalar diffraction theory (SDT). More specifically, the results presented here examine whether it is possible to create a constant color within a certain angular interval.

The overall aim is to explore the limits of structural colors. Applications for structural coloring include altering the color of certain products even though they are made of the same material by, e.g., changing the casting mold (plastic coloration, preprinted text) [10], new color appearances for product design, extremely durable colors due to UV stability [11], security labeling [12], improved optical materials (antireflective surfaces, such as moth eye structures [11], color filters [13], and solar cell films [14]), and more eco-friendly production processes that avoid the use of paint.

It is important to stress that structuring an objects surface is just one out of many fundamentally different methods of obtaining structural colors. Whereas gratings are often mass produceable due to cast molding and etching technologies, other methods with different purposes have been investigated, and a wide range of colors have been produced. Examples are surface plasmons on structured surfaces with deposited metallic layers [15,16], controlling absorption of coated metallic nanoparticles by controlling size parameters [17], and plasmon effects in carbon nanostructures [18]. All of these methods are capable of producing a variety of colors, but with little control of their appearance with respect to angular distribution. However, in optics the control of angular distribution and intensities for different purposes has been studied. In [19] intensities for periodic narrowband structures were optimized using the nongradient-based particle swarm optimization method. In [13] surfaces were optimized for solar cell use. A number of methods for improving antireflective surfaces exist, one example of which is given in [20]. The design of such gratings can be seen as a special case (periodic and single frequency) of the optimization formulation

developed in this work. To the best of the authors' knowledge the only other works exploring the possibilities of systematic design of surfaces with angular distributed colors are [21] and [22]. These papers consider more complex distributions of material based on the topology optimization approach (c.f. [23]).

The rest of the paper is organized as follows. Section 2 describes the implementation of the physics and how to obtain a far-field response for the reflection of a surface, Section 3 describes how to convert light spectra to RGB colors, Section 4 states the proposed optimization problem for color optimization, Section 5 presents results from the optimization, and Section 6 draws conclusions from the study.

2. SCALAR DIFFRACTION THEORY

All simulations in this paper are based on SDT. This theory is widely used for calculating diffraction gratings and surface scattering problems such as estimating reflection from rough surfaces. The simplicity of the mathematical expressions in SDT means that calculations can be made rapidly, and they are thus well suited for use in optimization problems where repetitive simulations are required. SDT—and its shortcomings—is well described in the literature, and the most central result in this theory is the relation between far-field irradiance, E, and a complex amplitude distribution, U, emerging from the surface of an aperture (which later will be used as a model for the actual surface profile) [24–26]:

$$E(x',y') = \frac{1}{\lambda^2 z^2}\left|\mathcal{F}\{U(x,y)\}\left(\frac{x'}{\lambda z},\frac{y'}{\lambda z}\right)\right|^2, \tag{1}$$

where λ is the wavelength and the coordinate components x, y, z, x', y' are as shown in Fig. 1. $\mathcal{F}$ represents the Fourier transform and is defined as

$$\mathcal{F}\{U(x,y)\}(\xi,\eta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(x,y)e^{-2\pi i(x\xi+y\eta)}\mathrm{d}x\mathrm{d}y. \tag{2}$$

The strength of SDT lies in the fact that many different physical setups can be described by defining U the right way. For the purpose of this paper, U will describe the complex phase modulation of a plane wave incident on a structured surface (see Fig. 2 for a 2D example). Using this setup for a material with constant reflection R, the complex amplitude distribution is then described as

$$U(x,y) = Ae^{2\pi i 2(h_{\mathrm{ref}}-h(x,y))/\lambda} = Ae^{-4\pi i h(x,y)/\lambda}e^{4\pi i h_{\mathrm{ref}}/\lambda}, \tag{3}$$

where $A = R$ in the region of interest and $A = 0$ (nonreflecting/totally absorbing) elsewhere. Since the phase information will be of no interest in the far field, the constant phase part with h_{ref} is dropped for convenience. Furthermore, this paper will not consider material properties, but just a full reflecting surface in free space. Therefore the amplitude will be dropped as well, so that the complex amplitude function is described as

$$U(x,y) = \mathrm{rect}_{A_s}(x,y)e^{-4\pi i h(x,y)/\lambda}, \tag{4}$$

where $\mathrm{rect}_{\Omega}(x)$ is a function that is 1 for $x \in \Omega$ and 0 otherwise. A_s is the region of interest.

A. Perception of Radiated Intensity

The expression in Eq. (1) is not easy to relate to the visual appearance of a surface. First, appearance is more naturally perceived with respect to viewing angle (in degrees), and second, irradiance is not proportional to the response of the human eye. Instead radiance will be used, as it possesses this property [27]. The coordinate system is first scaled and transformed such that

$$\hat{x} = x/\lambda, \qquad \hat{y} = y/\lambda, \qquad \hat{z} = z/\lambda \tag{5}$$

and

$$\alpha = \hat{x}/\hat{r}, \qquad \beta = \hat{y}/\hat{r}, \qquad \gamma = \hat{z}/\hat{r}, \tag{6}$$

where $\hat{r} = \sqrt{\hat{x}^2+\hat{y}^2}$. α and β are referred to as direction cosines, and their relation to a conventional spherical coordinate system is [28]

$$\alpha = \sin\theta\cos\phi, \tag{7}$$

$$\beta = \sin\theta\sin\phi, \tag{8}$$

$$\gamma = \cos\phi. \tag{9}$$

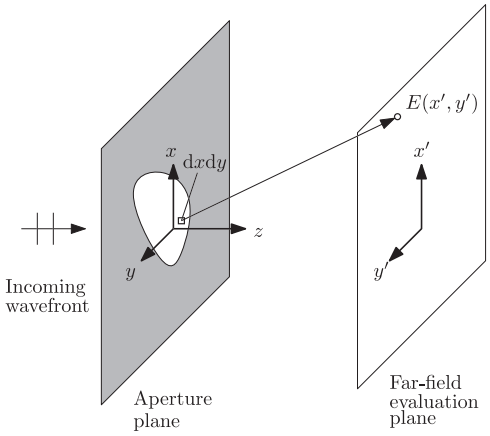


Fig. 1. Geometry for the scalar diffraction setup.

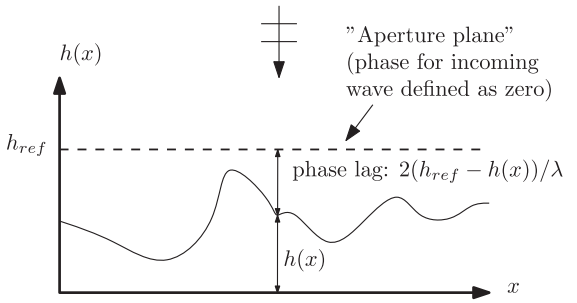


Fig. 2. Illustration of how a reflective grating can be turned equivalent to a complex distribution on an aperture plane. The phase lag is due to the extra distance traveled by the wave.

Using this transformation, the radiance L (and not irradiance) is given as [28,29]

$$L(\alpha,\beta) = \frac{\lambda^2}{A_s}|\mathcal{F}\{U(\hat{x},\hat{y},0)\}(\alpha,\beta)|^2 = \frac{1}{A_s}|\mathcal{F}\{U(x,y,0)\}(\alpha/\lambda,\beta/\lambda)|^2, \quad (10)$$

where A_s is the area of the aperture, and the last equality can be shown by substitution of variables in the Fourier integral. In spherical coordinates this means that the relation between radiance and irradiance far from the observer is given by $L(\theta,\phi) = r^2E(\theta,\phi)/(A_s\cos\theta)$, where r is the distance from the source to the observation point. In two dimensions (the $\phi = 0$ plane) the expression in Eq. (10) simplifies to

$$L(\alpha) = \frac{1}{L_0}|\mathcal{F}\{U(x)\}(\alpha/\lambda)|^2, \quad (11)$$

where L_0 is now the length of the aperture. This will be the starting point for all subsequent derivations.

B. Limitations of SDT

The advantage of SDT is the physical knowledge to be gained from its simplicity, and the disadvantage is that some physical properties are neglected due to its simplicity (see, e.g., [30] for a discussion on the interpretation of diffraction gratings). Inherent in the previous expressions is a small-angle (paraxial) approximation [28], meaning that the expressions lose their validity if energy ends up outside of the hemisphere above the surface. This is where $\theta \notin [-90^\circ, 90^\circ]$ (see, e.g., Fig. 5). This can in our case be corrected for [28] by normalizing L such that the energy when integrating over $\theta \in [-90^\circ, 90^\circ]$ corresponds to the incoming energy, since we assume full reflection. The extension is called nonparaxial SDT. Since we are optimizing for small angles and normal incoming light and do not want structures with features approaching the element resolution size of the model, which would give rise to this effect due to the high-spatial-frequency content [28], we have not found it necessary to do so—only when performing checks on the validity of the final result obtained. Other important limitations are that SDT takes only "first surface reflections" into account, which means that no multiple scattering or surface plasmon effects can be captured in this way, and that in two dimensions SDT best corresponds to TM-polarized light. This is because the TM polarization has the E-field going out of plane in Fig. 2, and the boundary conditions do not require zero field for electrical components parallel to a full reflecting plane. For components perpendicular to the plane, zero field is required, thus making TE polarization a worse approximation, since on nonhorizontal edges the boundary condition is violated.

C. Discretization of SDT

Normally, diffraction gratings described by SDT are evaluated using a fast Fourier transform (FFT). In this paper, another approach is proposed, which consists of representing the surface using rectangular sections that all have contributions that can be found analytically. This means that the formulation can be used for very large as well as small elements, always yielding a correct solution (to the SDT formulation) and not having to oversample or zero pad the FFT. Furthermore, the gradients for the system can be found analytically. It can also be implemented with a constant memory consumption independent of the number of elements. The trade-off for this is speed. If a continuous function must be analyzed, it can be approximated by taking the average continuous value in an interval, as when discretizing any other physical structure (see Fig. 3).

For all simulations in this paper, the rectangular sections all have a constant width d and are divided into N elements. If $\{h_n\}_{n=0}^{N-1}$ are the corresponding N heights, as seen in Fig. 3, then the height function is

$$h(x) \approx \sum_{n=0}^{N-1}\mathrm{rect}(x/d - n)h_n, \quad \text{where} \quad \mathrm{rect}(x) = \begin{cases} 1 & \text{for } 0 < x \le 1, \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

Finding the response of this structure is straightforward due to the linearity of the Fourier transform and the well-known solution of the Fourier transform of a rectangle:

$$\begin{aligned} L(\alpha) &= \frac{1}{L_0}|\mathcal{F}\{\mathrm{rect}_{L_0}(x)e^{-4\pi i h(x)/\lambda)}\}(\alpha/\lambda)|^2 \\ &= \frac{1}{Nd}\left|\mathcal{F}\left\{\sum_{n=0}^{N-1}\mathrm{rect}_{[0,1]}(x/d-n)e^{-4\pi i h_n/\lambda}\right\}(\alpha/\lambda)\right|^2 \\ &= \frac{1}{Nd}\left|\sum_{n=0}^{N-1}e^{-4\pi i h_n/\lambda}\mathcal{F}\{\mathrm{rect}_{[0,1]}(x/d-n)\}(\alpha/\lambda)\right|^2 \\ &= \frac{1}{Nd}\left|\sum_{n=0}^{N-1}e^{-4\pi i h_n/\lambda}de^{-\pi i d\alpha/\lambda}e^{2\pi i n d\alpha/\lambda}\mathrm{sinc}(d\alpha/\lambda)\right|^2 \\ &= \left|\underbrace{e^{-\pi i d\alpha/\lambda}\mathrm{sinc}(d\alpha/\lambda)}_{\text{const.shape}}\frac{1}{N}\sum_{n=0}^{N-1}\underbrace{e^{-4\pi i h_n/\lambda}}_{\text{height}}\underbrace{(e^{2\pi i d\alpha/\lambda})^n}_{\text{translation}}\right|^2. \end{aligned} \quad (13)$$

The different products of this expression all have a physical interpretation. First, the response is shaped by $C(\alpha) = e^{-\pi i d\alpha/\lambda}\,\mathrm{sinc}(d\alpha/\lambda)$, which is due to the choice of rectangles as the interpolation function. If d is large, then the sinc function will give rise to unavoidable zeros in the angular domain (this is also in agreement with the physics), but for $d\alpha/\lambda \to 0$, the expression converges toward 1, thus not influencing the response if the approximation of h is well resolved. $H_n = e^{-4\pi i h_n/\lambda}$ describes the phase change to which the height at n gives rise, and $T_n(\alpha) = (e^{2\pi i d\alpha/\lambda})^n$ modulates the phase change with a factor corresponding to the translation of the structure with respect to the origin. Note that $T_n = (T_1)^n$. It

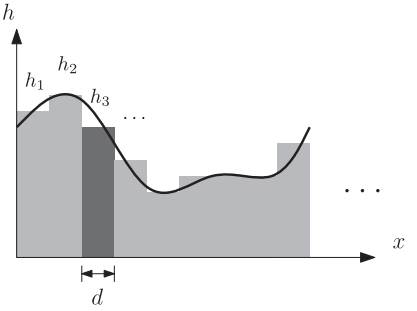


Fig. 3. Discretization of a surface into rectangular bumps.

can furthermore be seen from this expression how all element contributions are independent of each other. This means that there is no coupling or other interaction between the elements, which is also an implicit assumption when using SDT. The expression can now more compactly be written as

$$L(\alpha) = \left| C(\alpha) \frac{1}{N} \sum_{n=0}^{N-1} H_n(\alpha) T_n(\alpha) \right|^2. \tag{14}$$

1. Implementation

To calculate the responses in an efficient way, two steps are taken. First, H_n is precalculated and stored in matrices for a wide range of equally spaced heights, so when a height is given, the response is calculated as a linear interpolation of the two closest heights. This could in general give rise to two problems: first, such interpolation makes the gradient piecewise linear, thus not giving continuous derivatives for the gradient-based optimization algorithm, and second, a linear interpolation of numbers on the complex unit circle will not yield correct values. In practice, though, no problems have been observed regarding these issues, as the interpolation points are quite close. For all of the experiments in this paper, a spacing of 1 nm is used. In the second step, it is not necessary to calculate all the translation matrices, because of their recursive relationship. For our choice of implementation it is favorable with respect to speed to implement the calculations using a recursive approach defined by

$$P = C \frac{1}{N} \hat{P}_N, \quad \text{where} \quad \hat{P}_n = \begin{cases} T_1 \hat{P}_{n-1} + H_{N-1-n} & \text{for } n > 0, \\ H_{N-1} & \text{for } n = 0, \end{cases} \tag{15}$$

where P satisfies $L = |P|^2$. This formula can be intuitively understood as starting with the rightmost rectangle at the origin, translating it d, adding the element that should be to the left of it, translating the new structure d, and then continuing this process until all elements have been "pushed" into the right order. It is interesting to note that the T_1 and $\hat{P}_n$ are matrices whose size is determined as the product of the number of wavelengths and the number of reflection angles. The H_n values depend on the number of *height discretization* steps and the number of wavelengths, and there is as such no dependence on the number of elements used. The calculations therefore have a constant memory consumption independent of the number of elements—except the array in which the element heights are stored. The solving time increases linearly due to the recursive summation. This is because there is no coupling between the elements.

2. Gradient Calculation

As the optimization is gradient based, the gradients of the reflected field with respect to height changes are needed. Due to the fact that every element is decoupled from every other, the gradients with respect to the height variation that are required for the optimization can simply be found as

$$\begin{aligned} \frac{\partial P}{\partial h_j} &= C \frac{1}{N} \sum_{n=0}^{N-1} T_n \frac{\partial}{\partial h_j} e^{-4\pi i h_n/\lambda} \\ &= -\frac{4\pi i}{N\lambda} C T_j e^{-4\pi i h_j/\lambda} = -\frac{4\pi i}{N\lambda} C T_j H_j. \end{aligned} \tag{16}$$

3. COLOR CONVERSION

The color conversion is performed using

$$\mathrm{R} = \int_0^\infty I(\lambda) \bar{r}(\lambda) \mathrm{d}\lambda, \tag{17}$$

$$\mathrm{G} = \int_0^\infty I(\lambda) \bar{g}(\lambda) \mathrm{d}\lambda, \tag{18}$$

$$\mathrm{B} = \int_0^\infty I(\lambda) \bar{b}(\lambda) \mathrm{d}\lambda, \tag{19}$$

where $\bar{r}$, $\bar{g}$, $\bar{b}$ are the color-matching functions shown in Fig. 4 and $I(\lambda)$ is the light spectrum to be converted [31]. The color-matching functions are obtained by taking the CIE 1931 2° standard observer color-matching functions, $\bar{x}$, $\bar{y}$, $\bar{z}$, which are found in [32], and converting them to RGB color-matching functions by a linear transform with values as specified in [33], where the white reference used is the D65 illuminant that resembles the sun. This procedure is the same as in [22]. To simplify the notation, we will refer to colors as three-dimensional vectors, i.e., $\boldsymbol{C} = (\mathrm{R}, \mathrm{G}, \mathrm{B})^T$.

A. Randomization

In the literature it has been shown how a random height variation of a repeated base structure will spread the color response to a wider angular range than if no randomization is introduced, where grating modes will be visible. This has in particular been studied for the Morpho butterfly, and numerical as well as experimental results can be found in [6–9]. For this paper we assume that the same kind of randomization can be added to the developed designs, thus making them usable for coloring a surface without strong diffraction effects. Performing the actual randomization is beyond the scope of this paper.

4. OPTIMIZATION PROBLEM

The purpose of optimizing the surface structures is to be able to obtain a prescribed angular color response with as strong an intensity as possible. To do so, we have implemented an algorithm for working on the following optimization problem:

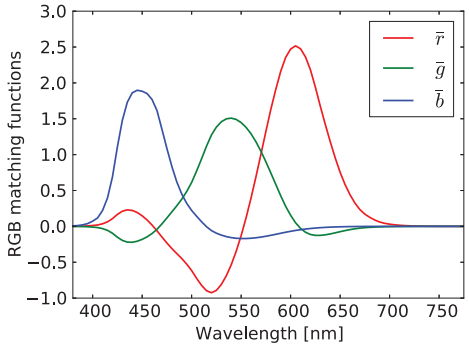


Fig. 4. CIE 1931 2° color-matching functions converted to RGB weighting functions with a D65 illuminant as reference.

$$\min_{\boldsymbol{h}} \; f(\boldsymbol{h}) = \max_{i=1,2,\ldots,M} \frac{-\|\boldsymbol{C}(\mathbf{h},\theta_i)\|^2}{\|\boldsymbol{C}_0(\theta_i)\|^2} + c_1 \frac{\|\nabla \boldsymbol{h}\|^2}{(Nd)^2},$$

$$\text{s.t.} \;\; \boldsymbol{h}_{\min} \le \boldsymbol{h} \le \boldsymbol{h}_{\max},$$

$$f_i(\boldsymbol{h}) = \frac{\|\boldsymbol{C}(\boldsymbol{h},\theta_i) \times \boldsymbol{C}_0(\theta_i)\|^2}{\varepsilon^2 \|\boldsymbol{C}_0(\theta_i)\|^2} - 1 \le 0, \tag{20}$$

where f is the objective function, the f_i denotes the M constraint functions, $\|\cdot\|$ is the Euclidean norm, $\boldsymbol{h} = (h_0, h_1, \ldots h_{N-1})$ is the design vector describing the heights of the profile, $\nabla \boldsymbol{h} = (h_1 - h_0, h_2 - h_1, \ldots h_{N-1} - h_{N-2})$ is a measure of the height differences between each element, $\boldsymbol{C}_0(\theta_i)$ is the prescribed color vector we want to obtain at the angle θ_i, M is the number of angles we consider in the optimization problem, $\boldsymbol{h}_{\min}$, $\boldsymbol{h}_{\max}$ are box constraints, and c_1 and ε are variables used to control the optimization problem, and they are described below.

The objective is formulated to maximize the worst-case (with respect to angle) desired color energy, subject to a constraint restricting energy in undesired color directions. This can be seen by considering

$$\begin{aligned} \|\boldsymbol{C} \times \boldsymbol{C}_0\|^2 &= \|\boldsymbol{n} \sin\phi \|\boldsymbol{C}\| \cdot \|\boldsymbol{C}_0\|\|^2 \\ &= \|\boldsymbol{C}\|^2 \cdot \|\boldsymbol{C}_0\|^2 \cdot \sin^2\phi, \end{aligned} \tag{21}$$

where ϕ is the angle between the two three-dimensional color vectors. Here it is seen how the squared cross product is proportional to the product of $\sin^2\phi$ and its own length squared. This means that the term approaches zero for vectors in the same direction or vectors with no intensity. Since we seek to maximize the color intensity, the algorithm will in practice try to satisfy the constraint by having a small angle between the vectors and not by keeping the intensity low. The variable ϵ is then used to control how small the angle should be without violating the constraint and has been set to $\epsilon = 0.01$ for all optimizations in this paper. This ensures that we obtain a color close to the prescribed color. The parameter c_1 is used for regularization by weighting the total variation of the design and is discussed in detail in the next section.

Since the optimization problems considered all require a symmetric color response, we have imposed a symmetry condition on the physical design.

The problem has been solved using the method of moving asymptotes [34]. This is a gradient-based optimization algorithm, and the gradient of the objective and the constraints can be calculated analytically using the chain rule, the linear color transform, and Eq. (16). Other nonlinear solvers handling inequality constraints could have been used as well.

A. Regularization Scheme

Various regularization techniques reviewed in [35] have been tried out, but based on experiments we prefer a penalization of the 2-norm of the design gradients, as proposed for other problems in [36]. The implementation of this is what can be seen as the last term in the objective function of Eq. (20). By doing so, a penalty for having huge jumps between neighboring elements is introduced, as well as a penalty for having sharp edges compared to soft edges. This is because the height differences are squared, and hence jumps between 0 and 1 are much more heavily penalized than those between 0 and 0.5—e.g., consider how three elements defined such that $\boldsymbol{h} = (0, 0, 1)$ will be penalized more than $\boldsymbol{h} = (0, 0.5, 1)$. The obvious goal of removing small features is therefore present, and furthermore the preference for soft curves makes it easier for the optimizer to move elements from one level to another. This is important because we have found that the cost function often does not have monotonic behavior in going from one level to another, causing the (gradient-based) optimizer to maintain its "levels" when they are found. At the same time, no minimum length scale is imposed by this method, leaving the optimization algorithm free to create features of any desired size.

In order to avoid any bias in the results caused by the regularization, a continuation approach is used. In contrast to normal continuation approaches, where an optimization is initialized with a large penalty/filter parameter and then it is gradually decreased, here the design is started with c_1 from Eq. (20) set to zero, and then gradually c_1 is increased when the optimizer approaches a convergent design (for these specific cases, a Karush–Kuhn–Tucker norm smaller than 0.1 has been used as the criterion) to a large value of c_1 (between 30 and 40) and afterward gradually decreased—using the same scheme—until $c_1 = 0$ and the optimizer is then continued until convergence. Due to the color constraint in the formulation of the optimization problem, we retain a design with the desired color, no matter how large we set c_1, and the maximum c_1 value is therefore a parameter defining how many features we are willing to remove. The maximum size of c_1 is found not to be very critical for the overall features and performance of the final design.

B. Setup

To obtain the following results, the setup is as shown in Fig. 5. That is, the optimization is performed for a 2 μm wide surface with light incident along the normal of the surface, which is the only contribution to the far field. The surroundings are considered absorbing, not present or in other ways nonreflecting. The small surface size is chosen to restrict the problem to a size that is easier for the optimization to handle, and because, if a surface is to be colored this way, the most feasible and production-friendly way would seem to be to find a small structure that can then be repeated with added randomness [9]. Another issue is that for larger structures, the incoherence of light would have to be taken into account [8]. The optimization is carried out for 20 equidistant wavelengths between 400 and 700 nm. It is found that fewer wavelengths make it difficult to resolve especially green colors well. The spacing between angles specified for optimization is 1°, and the element spacing, d, is set to 10 nm.

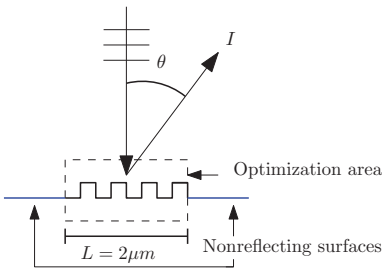


Fig. 5. Physics for the optimization problem.

5. RESULTS

The optimization is performed for two different setups: one for a simple specular reflection, and another for a constant color in an angular range of -10° to 10°. The first setup is used to prove the validity of the procedure by comparing it to analytical calculations—and also to reveal certain features of the optimized designs—whereas the second setup tests its usefulness for designing more complex color effects that cannot be predicted by intuition or simple formulas.

It should be noted that since the level of reflected energy is determined *a priori* by the SDT formulation (e.g., no extra absorption or transmission effects possible, such as surface plasmons), then the colors suppressed in a certain angular interval will be present outside this interval. Roughly speaking, the complementary color to $\boldsymbol{C}_0$ will be spread out in the rest of the angular domain. This means that a structure scattering blue around the specular direction will have the red and green color components "smeared" out in the rest of the angular domain, as seen in Fig. 6. Furthermore, all color plots in the following take the form seen in Fig. 6, with curves for R, G, and B on a background colored by the combined RGB value of the curves to illustrate the actual color.

A. Designing Specular Reflection

As a benchmark case for the proposed optimization procedure, optimization of the specular reflection for light incident normal to a structure is used. For transmission gratings (by scaling the height, the results can be used for transmission gratings as well), this type of structure is referred to as zero-order diffraction (ZOD) and finds its use in, e.g., color filters.

1. Analytical Solution for the Binary Grating

Using SDT, the expression for zero-order reflection simplifies a great deal, since this corresponds to putting $\alpha = 0$ in Eq. (13):

$$L(0) = \left| \frac{1}{N} \sum_{n=0}^{N-1} e^{-4\pi i h_n/\lambda} \right|^2. \tag{22}$$

Note how the elements lose their dependence on position, since the translation product becomes unity. This is due to the fact that we consider light incident normal to the surface and because of the simple formulation of SDT. In practice the positioning of the elements can be of importance, since the geometry should not violate the basic assumptions of SDT. This means that, e.g., a structure with only one big trench will be better than one with many small trenches, since the electromagnetic boundary condition on the vertical edges are ignored by SDT and the geometry will contain more high-spatial-frequency content [28].

A particularly simple and relevant case is when we restrict the heights to take one of two values. This is known as a binary phase grating and is easy to produce, e.g., by using etching processes, and is therefore widely used. For this, assume that we allow only two values for h_n, which can be defined such that h_1 is the height difference between the two elements and $h_0 = 0$ [see Fig. 7(c)]; then the response at $\theta_{\text{out}} = 0^{\circ}$ will be expressed by the following simple analytical expression:

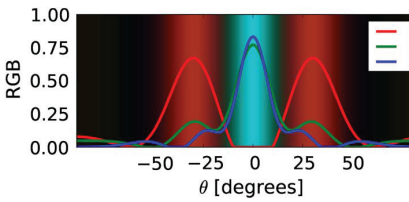


Fig. 6. Reflection for a full reflective surface to illustrate how the color not scattered in the desired directions will be scattered in other directions. In this plot blue and green are scattered close to the specular direction, where red is suppressed, meaning that red will have to be scattered somewhere else.

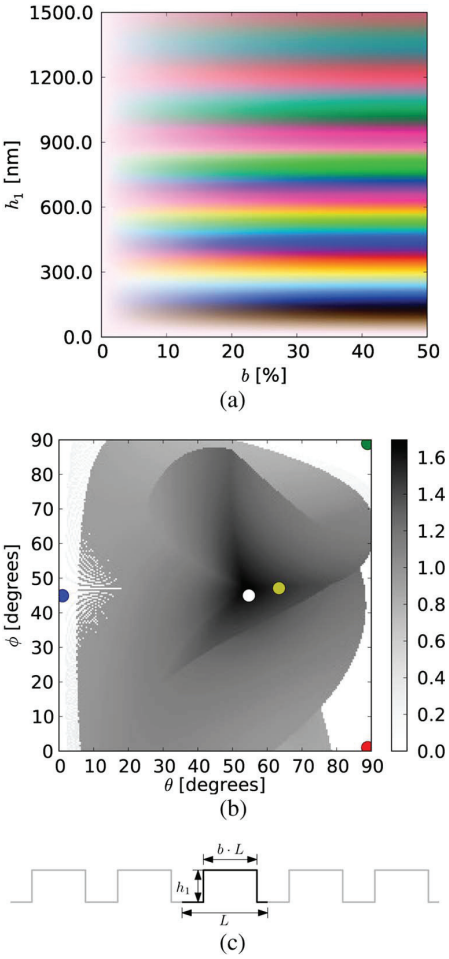


Fig. 7. Analysis of possible colors produceable from the specular mode of a binary phase grating. The maximum height has been limited to 1500 nm. (a) Map of binary phase grating zero-order colors as a function of height and duty cycle (a few RGB values have been cut, as they were negative or exceeded 1) and (b) possible "color directions" for the colors in (a), where the color vector has been transformed into polar coordinates and the gray scale indicates the (largest obtained) value of the length, r, of the color vector. The red, green, and blue points have been indicated, as well as white and black (which overlap, since they only vary in intensity), which are represented by a white circle, and the obtainable test reference is indicated with a yellow circle. (c) Example of how a structure can be realized for certain parameters of h_1 and b.

$$L(b,h_1) = \left| \frac{\sum_{n\in H_0} e^{-4\pi i h_0/\lambda}}{\sum_{n\in H_0} 1} + \frac{\sum_{n\in H_1} e^{-4\pi i h_1/\lambda}}{\sum_{n\in H_1} 1} \right|^2$$
$$= |(1-b) + be^{-4\pi i h_1/\lambda}|^2, \quad (23)$$

where b is the duty cycle (or filling fraction) of h_1. Since whether or not h_1 is defined as a positive or negative quantity with respect to h_0 (making h_0 the bump if h_1 is negative) gives rise to a change only in the far-field phase—and not the amplitude—the function $L(b,h_1)$ for $b \in [0,1]$ is symmetric around $b = 0.5$, and we therefore need to consider only the interval $b \in [0,0.5]$. By converting the spectra obtained with different choices of h_1 and b to colors, it is possible to map the space of realizable colors for this setup. The result can be seen in Fig. 7(a). Furthermore, converting the set of color vectors from Cartesian coordinates to spherical coordinates using $r = \sqrt{R^2+G^2+B^2}$, $\theta = \arccos B/r$, $\phi = \arctan G/R$, we can plot the possible color directions for binary gratings to obtain an idea of how well we can fill the color space using the specular mode of binary gratings. This mapping procedure makes the color a two-dimensional quantity, which is similar to plotting x and y from a CIE xyY color space. The result is shown in Fig. 7(b). From these plots it can be seen that not all colors are realizable using this approach [as indicated in Fig. 7(b); e.g., pure red, green, and blue cannot be achieved]. To verify the optimization algorithm for more complex designs, a realizable color using a binary grating and red, green, and blue will be the targets. As a more or less random choice, the yellowish color for $h_0 = 275$ nm, $b = 25\%$ was chosen as a reference color for the optimization, which gives rise to $(R,G,B) = (0.89,0.96,0.65)$, or $(r,\theta,\phi) = (1.46, 63.4^\circ, 47.1^\circ)$ for specular reflection. In addition to the pure red, green, and blue color, this was used as the optimization goal $\boldsymbol{C}_0$.

2. Detailed Analysis

Going into more detail with the color space, it is possible to find the same yellowish response (generated by a different spectrum) at $(h_0,b) \approx (550\text{ nm}, 11\%)$, and so it should be noted that two binary solutions exist. If we instead consider the pure red, green, and blue color and try to find the points in the color space matching them best—that is, giving the minimum valid value of ε in Eq. (20)—we find, respectively, that $(h_0,b) \approx (121\text{ nm}, 50\%)$ with $\varepsilon \approx 0.02$, $(h_0,b) \approx (129\text{ nm}, 50\%)$ with $\epsilon \approx 0.10$, and $(h_0,b) \approx (146\text{ nm}, 50\%)$ with $\epsilon \approx 8.3 \cdot 10^{-4}$. Looking at Fig. 7(a), the colors look very dark in that area, which means that the low intensity is helping satisfying the constraint. These solutions are not what the optimization algorithm will aim for, since they have extremely low intensities and will therefore not satisfy the parallel requirement in Eq. (20). If we exclude heights below 200 nm and find the best match for heights above that, we get, respectively, $(h_0,b) \approx (365\text{ nm}, 50\%)$ with $\epsilon \approx 0.23$, $(h_0,b) \approx (790\text{ nm}, 50\%)$ with $\epsilon \approx 0.16$, and $(h_0,b) \approx (432\text{ nm}, 50\%)$ with $\epsilon \approx 0.09$. This study confirms the results in Fig. 7(b), where it is seen that even though a clear blue does not exist in the color space of simple binary gratings, we can find solutions close to it. Furthermore, it is interesting to see that green needs a fairly large difference in height. This can probably be explained by the fact that green is in the middle of the spectrum, and therefore both the lower and upper parts of the spectrum need to be suppressed, whereas red and blue each occupy their part of the spectrum and we just need to suppress the rest.

3. Optimization without Regularization

Running the optimization formulated in Eq. (20) for $\boldsymbol{C}_0 = (0,0,1)$ (blue) at $\theta_0 = 0^\circ$, with $\epsilon = 10^{-2}$, no regularization ($c_1 = 0$), and an initial guess with random heights varying from 250 to 1250 nm, gives the result obtained in Fig. 8. Note that not starting from zero leaves space for the optimization to decrease the heights if necessary. The structure is seen to have very fine one-element features. This is undesirable, as the small features scatter energy in evanescent modes, and paraxial SDT does not take this into account, leading to nonphysical solutions. Besides this, there will be multiple reflections/scattering and internal scattering that are not taken into account, and the result will probably be very different for TE polarization due to the small features. We therefore obtain structures that are hardly realizable when the regularization term is not included. Analyzing the structure using nonparaxial SDT, the color response changes significantly, as seen in Fig. 8, and it is clear that the design will not satisfy our original objective.

To solve all of these problems at once, regularization is used—as discussed in Section 4.A—to improve the quality of design with respect to the length scale.

4. Optimization with Regularization

Using the regularization scheme from Section 3.A and rerunning the optimization, the results presented in Fig. 9 are obtained for blue as well as for the other three prescribed colors. In the figure, $|f|$ indicates the final objective value and N_{it} the number of iterations used. The number of

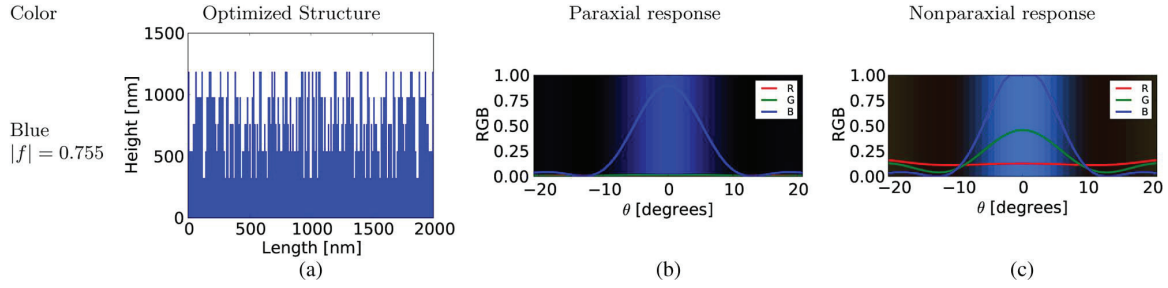


Fig. 8. Optimization of ZOD for blue color without any regularization. (a) The obtained design, (b) the color response seen by the optimization algorithm, and (c) an analysis of the final structure using nonparaxial SDT. This shows that SDT does not capture the physics well, due to scattering from small feature sizes.

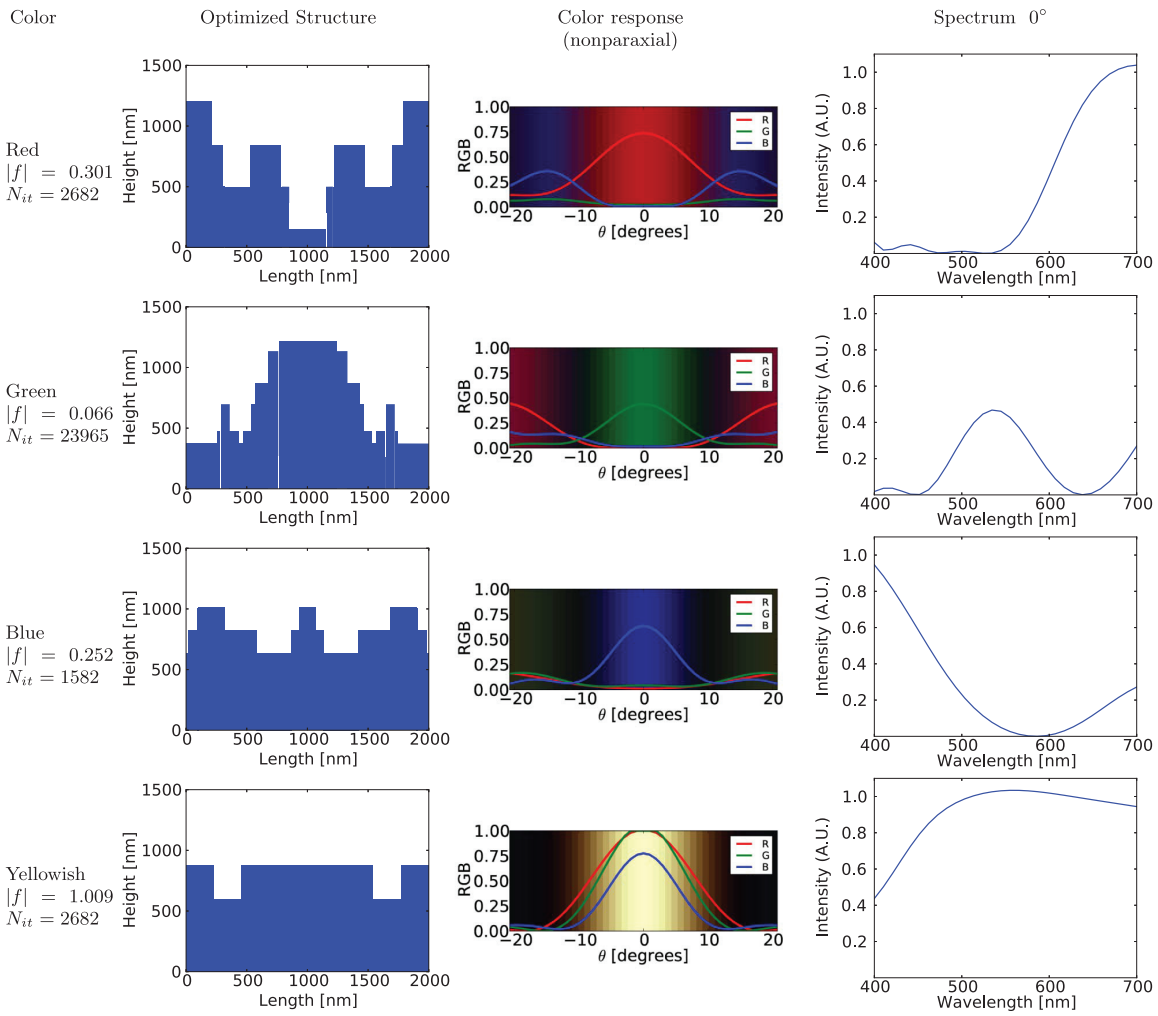


Fig. 9. Optimization of ZOD using regularization for different colors. All parameters except $\boldsymbol{C}_0$ and c_1 are unchanged from the previous example. The color coordinates for the prescribed design colors are (1, 0, 0), (0, 1, 0), (0, 0, 1), and (0.89, 0.96, 0.65), respectively.

iterations are for most designs a couple thousand, which is reasonable, since design evaluation is fast. For more computationally heavy optimizations, iterations may be decreased by tuning the regularization parameters. The final results were generated using nonparaxial SDT to show that the results are still satisfactory. A much simpler structure for blue is now obtained at the expense of a decrease in objective value. The structure is divided into three levels, and no single element features are present. From Fig. 7(b) it may be seen how one level would not be enough to generate blue. For red, the same features may be observed, using only four levels. However, for green, it seems that many levels are needed to obtain a good result, but the number of iterations also indicates that the algorithm has had difficulties obtaining a good result. An explanation for this could be that the green color is in the middle of the spectrum, and the grating therefore should act more as a bandpass filter, while red and blue each occupy their end of the visible spectrum, meaning that their behavior resembles more that of a low- or high-pass filter, which is normally easier to realize. For the test color, it may be seen that the optimizer found a binary solution (without regularization, a multilevel structure with worse performance was found). Its height and duty cycle are $h_1 = 278$ nm, $b = 23\%$, which is close to the reference $h_1 = 275$ nm, $b = 25\%$. The small deviations are probably due to the fact that this solution still satisfies the color constraint but obtains a better objective value.

Running the optimization for the yellowish color with the starting guess being $(h_0, b) \approx (550\text{ nm}, 11\%)$, which is shown in Section 5.A.2 to also be a valid solution, we end up with a slightly perturbed version of the starting guess and an objective of $|f| = 0.946$, which is slightly lower than what is observed in Fig. 9. This confirms how multiple solutions might exist. In practice several initial guesses have been tried for each color, but have ended up having more or less the same efficiency.

In general the optimization is found to work well, and a wide range of ZOD designs can be created using this method. The designs all have large flat areas divided into levels, even though there is nothing in the formulation preferring that kind of design (either with or without regularization), and from this

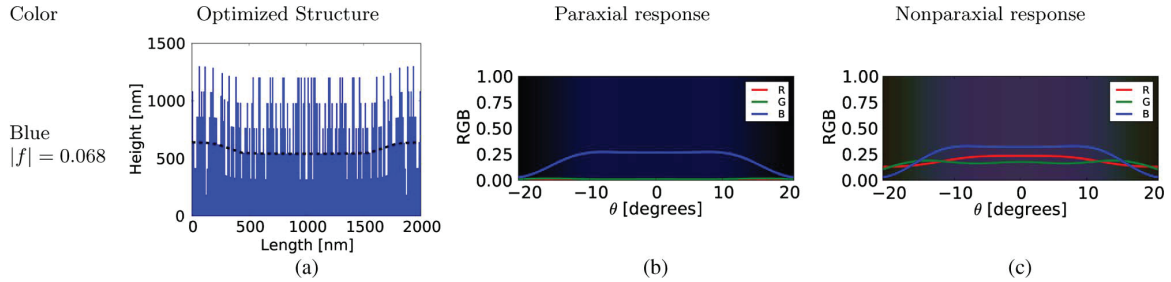


Fig. 10. Optimization for blue for $\theta_{\text{out}} \in [-10^{\circ}, 10^{\circ}]$ with no regularization (all other parameters are the same as for the ZOD example). (a) The design obtained, with the dashed line indicating one of the envelopes seen in the structure, (b) the color response seen by the optimization algorithm, and (c) an analysis of the final structure using nonparaxial SDT.

it may be concluded that at least for the colors considered in the examples (and a number of other colors tried by the authors), a few levels are enough to tune a surface to a certain color in the specular mode. It is also worth noting that we have observed that for some colors the regularization method—in addition to yielding a feasible solution—also leads to a better objective value than without regularization.

B. Wide-Angle Color Optimization

Optimizing for wide-angle color response without regularization encounters the same problems as for the ZOD case: performing the same optimization run—now using an angular interval of $\theta_{\text{out}} \in [-10^{\circ}, 10^{\circ}]$—yields the result shown in Fig. 10. Again, SDT did not resolve the color very well, and the structure has one-element features, so regularization is again required.

One interesting feature to take note of in Fig. 10 is that it resembles the result in Fig. 8, but with a nonconstant envelope for each level. For the colors where the nonregularized ZOD optimization results in only a few flat levels of height difference, the wide-angle optimization often turns out to be at the same levels, but with a nonconstant envelope. Extracting the envelope and simulating this alone shows that the envelope shape has the effect of scattering $\boldsymbol{C}_0$ in a wider angle than a flat surface would. This is illustrated in Fig. 11, where one of the envelopes from Fig. 10 was extracted and a much flatter response may be seen in the blue region compared to a flat surface. This indicates that some designs can be obtained when choosing a color by specifying a height difference, and then finding an appropriate envelope shape of these heights to spread out the color in the specified angular domain.

1. *Designs with Regularization*

To avoid small design features such as those in Fig. 10, we introduce regularization using the same scheme as in Section 4.A. The results are shown in Fig. 12. What is interesting to note here is that trends similar to the ones from the ZOD optimization can be found: the structure of the test color and blue are the simplest, having only three levels and some slight curvature within more or less the same height difference as for the ZOD result. For green the profile extends to a slightly larger height range but otherwise shows the same trends as blue. For red the optimized structure seems rather simple, but the response is not flat. This means that when the lowest

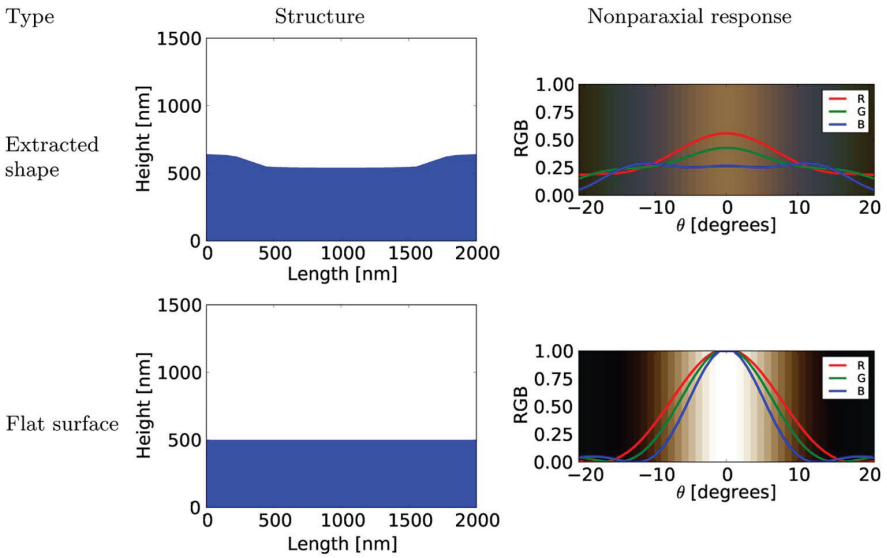


Fig. 11. Comparison of the envelope indicated by a dashed line in Fig. 10(a) and a flat surface having the same length. It is clearly seen how the shape scatters blue in an almost flat interval from -10° to 10° compared to the flat response.

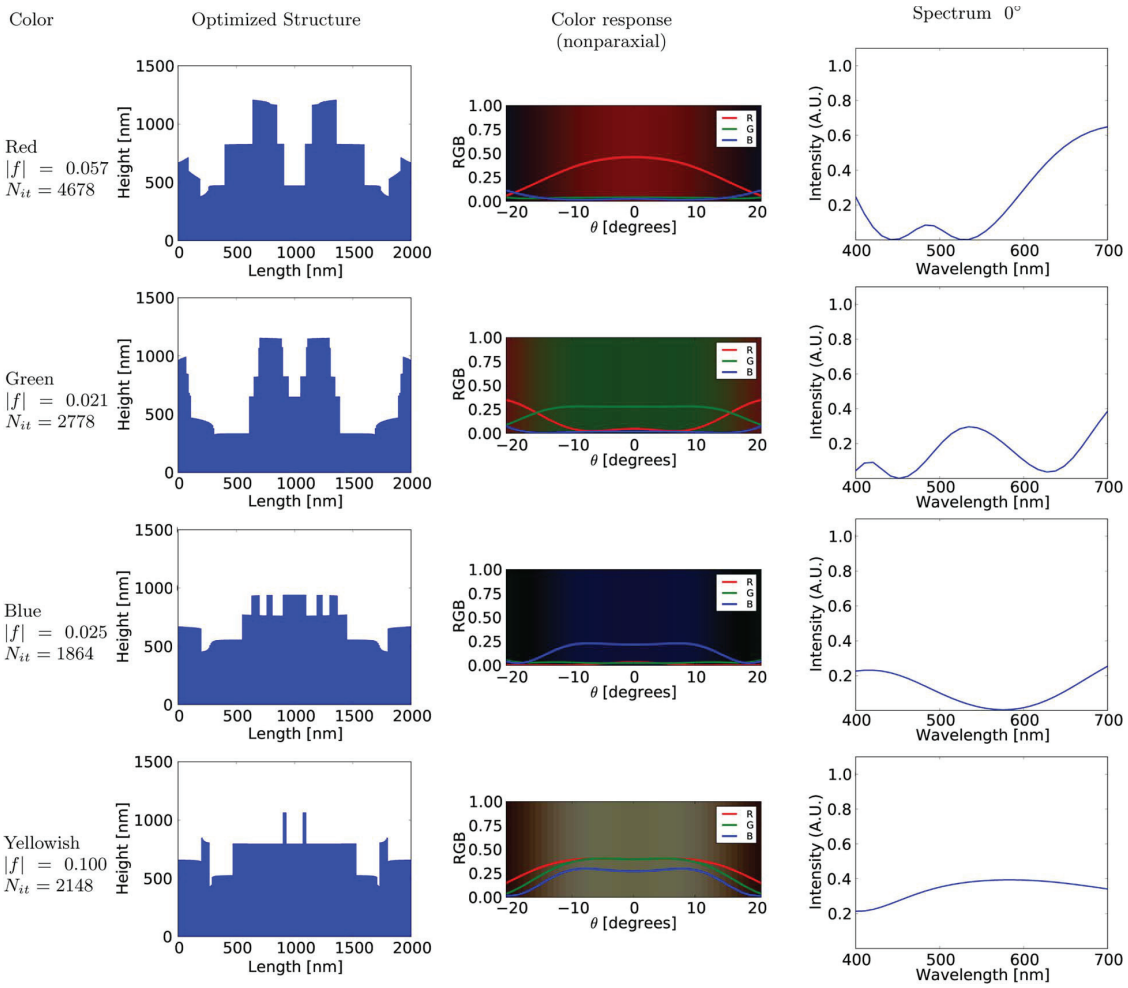


Fig. 12. Optimization for different colors for $\theta_{\text{out}} \in [-10^{\circ}, 10^{\circ}]$ using regularization. Except for $\boldsymbol{C}_0$ and c_1, all the parameters are unchanged from the previous example. The color coordinates for the prescribed design colors are (1, 0, 0), (0, 1, 0), (0, 0, 1), and (0.89, 0.96, 0.65), respectively.

intensities are raised to $\pm 10^{\circ}$, the other intensities are raised as well. A totally flat response for red was probably not obtained because red light has the longest wavelengths, and thus in general it will scatter less for the 2 μm surface chosen for these experiments. To verify this, we have tried the same procedure with $d = 20$ nm, giving a design domain of 4 μm, and the result obtained here is indeed flat.

In general it seems that the more complex the ZOD design is, the more complex the wide-angle design is, and that overall the procedure seems to be able to find structures satisfying the constraints independently of the choice of color.

6. CONCLUSIONS

In this paper we have shown how to formulate a structural color optimization problem and use it for the optimization of fully reflecting surface profiles to create on-demand angular color effects. The modeling is carried out using SDT discretized and solved in a novel way, but all kinds of electromagnetic methods can be used. The formulation is able to find structures for ZOD responses as well as wide-angle responses. For surface profiles it is observed that not all colors are equally easy to obtain, and explanations for this have been given. First, colors that can be created by the interference of two different heights/levels seem to be easier to obtain; second, since the width of the design domain was fixed and it diffracts to wider angles for lower wavelengths, blue colors seem easier to obtain for wide-angle responses; and third, colors that can be obtained by high-/low-pass filtering instead of bandpass/bandstop filtering seem easier to obtain—most likely because these types of filters in general are easier to synthesize. It was also observed that all designs contain sharp edges as a part of the final design, which is a consequence of the fact that strong constructive/destructive interference is needed to filter the colors of the visible spectrum.

As a part of the study, the colors produceable in the zero-order mode by two-level interference were also derived analytically and mapped in Fig. 7(a). To avoid mesh-dependent features in the design, a regularization scheme was introduced that exploits the fact that the color was imposed as a constraint to the optimization problem. This made it possible to obtain design with only the essential features, retaining a good response.

ACKNOWLEDGMENTS

This project was supported by the Danish National Technology Foundation through the projects ODAAS and NANOPlast, and by the Villum Foundation through the NextTop project.

REFERENCES

1. P. Vukusic and J. R. Sambles, "Photonic structures in biology," Nature **424**, 852–855 (2003).
2. S. Kinoshita and S. Yoshioka, "Structural colors in nature: the role of regularity and irregularity in the structure," ChemPhysChem **6**, 1442–1459 (2005).
3. A. Saito, "Material design and structural color inspired by biomimetic approach," Sci. Tech. Adv. Mater. **12**, 064709 (2011).
4. S. Kinoshita, *Structural Colors in the Realm of Nature* (World Scientific, 2008).
5. R. Hooke, Micrographia, 1665, http://www.gutenberg.org.
6. S. Kinoshita, S. Yoshioka, Y. Fujii, and N. Okamoto, "Photophysics of structural color in the Morpho butterflies," Forma **17**, 103–121 (2002).
7. R. T. Lee and G. S. Smith, "Detailed electromagnetic simulation for the structural color of butterfly wings," Appl. Opt. **48**, 4177–4190 (2009).
8. A. Saito, M. Yonezawa, J. Murase, S. Juodkazis, V. Mizeikis, M. Akai-Kasaya, and Y. Kuwahara, "Numerical analysis on the optical role of nano-randomness on the Morpho butterfly's scale," J. Nanosci. Nanotechnol. **11**, 2785–2792 (2011).
9. A. Saito, Y. Miyamura, Y. Ishikawa, J. Murase, M. Akai-Kasaya, and Y. Kuwahara, "Reproduction, mass-production and control of the Morpho-butterfly's blue," Proc. SPIE **7205**, 720506 (2009).
10. Front Page—Nanoplast, http://www.nanoplast.dk/, 2013.
11. O. Karthaus, *Biomimetics in Photonics* (CRC Press, 2012).
12. K. Yu, T. Fan, S. Lou, and D. Zhang, "Biomimetic optical materials: integration of natures design for manipulation of light," Prog. Mater. Sci. **58**, 825–873 (2013).
13. M. J. Uddin and R. Magnusson, "Highly efficient color filter array using resonant Si_3N_4 gratings," Opt. Express **21**, 12495–12506 (2013).
14. X. Sheng, S. G. Johnson, J. Michel, and L. C. Kimerling, "Optimization-based design of surface textures for thin-film Si solar cells," Opt. Express **19**, A841–A850 (2011).
15. K. Kumar, H. Duan, R. S. Hegde, S. C. W. Koh, J. N. Wei, and J. K. W. Yang, "Printing colour at the optical diffraction limit," Nat. Nanotechnol. **7**, 557–561 (2012).
16. Y. R. Wu, A. E. Hollowell, C. Zhang, and L. J. Guo, "Angle-insensitive structural colours based on metallic nanocavities and coloured pixels beyond the diffraction limit," Sci. Rep. **3**, 1194 (2013).
17. G. Park, C. Lee, D. Seo, and H. Song, "Full-color tuning of surface plasmon resonance by compositional variation of Au@Ag core-shell nanocubes with sulfides," Langmuir **28**, 9003–9009 (2012).
18. T. Xu, H. Shi, Y. Wu, A. F. Kaplan, J. G. Ok, and L. J. Guo, "Structural colors: from plasmonic to carbon nanostructures," Small **7**, 3128–3136 (2011).
19. Z. Shi and Y. Gao, "Design of Dammann gratings by particle swarm optimization," in *Symposium on Photonics and Optoelectronics*, June 2010.
20. X. Li, Q. Tan, and G. Jin, "Surface profile optimization of antireflection gratings for solar cells," Optik **122**, 2078–2082 (2011).
21. K. S. Friis and O. Sigmund, "Robust topology design of periodic grating surfaces," J. Opt. Soc. Am. B **29**, 2935–2943 (2012).
22. J. Andkjær, V. E. Johansen, and O. Sigmund, "Inverse design of nano-structured surfaces for color effects," J. Opt. Soc. Am. B **31**, 164–174 (2014).
23. J. S. Jensen and O. Sigmund, "Topology optimization for nano-photonics," Laser Photon. Rev. **5**, 308–321 (2011).
24. J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed., McGraw-Hill Physical and Quantum Electronics Series (McGraw-Hill, 1996).
25. J. E. Harvey, "Radiometry rocks," Proc. SPIE **8483**, 848304 (2012).
26. D. C. O'Shea, "Scalar diffraction theory," in *Diffractive Optics: Design, Fabrication, and Test* (SPIE, 2003), pp. 17–35.
27. P. Dutré, K. Bala, and P. Bekaert, *Advanced Global Illumination* (A K Peters/CRC Press, 2006).
28. J. E. Harvey, C. L. Vernold, A. Krywonos, and P. L. Thompson, "Diffracted radiance: a fundamental quantity in nonparaxial scalar diffraction theory," Appl. Opt. **38**, 6469–6481 (1999).
29. J. E. Harvey, C. L. Vernold, A. Krywonos, and P. L. Thompson, "Diffracted radiance: a fundamental quantity in nonparaxial scalar diffraction theory: errata," Appl. Opt. **39**, 6374–6375 (2000).
30. A. W. Lohmann, "About the philosophies of diffraction," in *International Trends in Optics*, J. W. Goodman, ed. (Academic, 1991), Chap. 11, pp. 155–164.
31. R. S. Berns, F. W. Billmeyer, and M. Saltzman, *Billmeyer and Saltzman's Principles of Color Technology* (Wiley-Interscience, 2000).
32. CIE: International Commission on Illumination, http://www.cie.co.at/.
33. International Electrotechnical Commission, "Colour management—default RGB colour space—sRGB," IEC 61966-2-1, 1999.
34. K. Svanberg, "The method of moving asymptotes—a new method for structural optimization," Internat. J. Numer. Methods Engrg. **24**, 359–373 (1987).
35. O. Sigmund, "Morphology-based black and white filters for topology optimization," Struct. Multidiscip. Optim. **33**, 401–424 (2007).
36. T. Borrvall, "Topology optimization of elastic continua using restriction," Arch. Comput. Methods Eng. **8**, 351–385 (2001).

Publication P2

Inverse design of nanostructured surfaces for color effects

Inverse design of nanostructured surfaces for color effects

Jacob Andkjær, Villads Egede Johansen, Kasper Storgaard Friis, and Ole Sigmund*

Technical University of Denmark, Department of Mechanical Engineering, 2800 Kgs. Lyngby, Denmark
**Corresponding author: sigmund@mek.dtu.dk*

Received May 28, 2013; revised November 27, 2013; accepted November 27, 2013;
posted December 2, 2013 (Doc. ID 191305); published December 24, 2013

We propose an inverse design methodology for systematic design of nanostructured surfaces for color effects. The methodology is based on a 2D topology optimization formulation based on frequency-domain finite element simulations for E and/or H polarized waves. The goal of the optimization is to maximize color intensity in prescribed direction(s) for a prescribed color (RGB) vector. Results indicate that nanostructured surfaces with any desirable color vector can be generated; that complex structures can generate more intense colors than simple layerings; that angle independent colorings can be obtained at the cost of reduced intensity; and that performance and optimized surface topologies are relatively independent on light polarization.

OCIS codes: (230.1950) Diffraction gratings; (230.4170) Multilayers; (310.6628) Subwavelength structures, nanostructures; (330.1690) Color; (330.7326) Visual optics, modeling.
http://dx.doi.org/10.1364/JOSAB.31.000164

1. INTRODUCTION

Structural colors are caused by interference effects from light interaction with nanostructures rather than by pigments. The vivid colors generated by the nanostructures are evident in many animals and plants [1], e.g., the *Morpho* butterfly displays a blue color due to a tree-like multilayer structure [2] and the *Papilio palinurus* displays a bright green color [3]. Even though structural color effects have been of scientific interest for centuries [4], renewed interest has been sparked in recent years due to advances in nanotechnology.

Particularly the wings of the *Morpho* butterfly, as already mentioned, have been studied intensively for their optical properties and structural coloration. With the development of the electron microscope, it has been possible to observe the periodic ridge structure giving rise to the structural color. The ridge structure has a treelike shape that resembles a multilayer structure. The multilayer structure is tuned such that a reflection band occurs for the "blue" wavelengths. However, the color appears blue in a too-wide angular range compared to traditional grating and multilayer theories. The explanation is the height randomization of individual ridges, which destroys interference and hence causes smooth angular color appearance. In [5] a numerical analysis is conducted on the 3D nanostructure, which is the source of the blue color of the *Morpho* butterfly wings. Further studies on numerical modeling of *Morpho* structures and their randomness can be found in [6], and a methodology for rendering of *Morpho* wings can be found in [7].

Much research has been centered on analyzing and artificially replicating the structures found in nature—so-called biomimetics or biomimicry [3,8]. However, in the present work we take another approach. Instead of mimicking what nature produced, we define the inverse design problem: find a dielectric nanostructure that displays a prescribed color effect. In this way we are not limited to what can be found in nature—we should be able to produce any desired color effect. The work is motivated by the developments in nanotechnology, improving the possibilities for achieving submicrometer features on surfaces also in large batch sizes. Examples of future applications are within design of nondegradable colors [8] and paint-free productions [9].

In previous studies [10–12], a method based on topology optimization [13] has been formulated for designing nanostructured surfaces with extreme reflection or transmission properties. Here we extend [12] by including a near- to far-field transformation of the reflected light and converting the scattered far-field spectrum to colors. With this extension, we obtain a methodology based on topology optimization for designing nanostructured multilayered surfaces displaying prescribed structural color properties.

Topology optimization is an inverse design methodology based on repeated analyses (finite element or finite difference analysis in frequency or time domain), gradient computations and material redistribution based on mathematical programming concepts. Topology optimization was originally developed for mechanical design problems [13,14] but has since then been extended to a number of other physics areas, including nanophotonics [15–17], antennas [18], and metamaterials [19].

The power of the topology optimization method is that it can suggest novel geometries without any geometrical restrictions. Obviously, this extreme design freedom may also result in structures that are difficult to manufacture using currently available technologies. If the manufacturing technology and its limitations are known, one may introduce manufacturability as a part of the optimization problems as, e.g., seen in our previous work, where we introduced a connectivity constraint [11]. Obviously, imposing constraints on the geometrical freedom limits the achievable color response. To give the reader an idea of this aspect, we compare the performance of the

designs achieved with full geometrical freedom with corresponding performances for designs that are limited to be simple layerings.

The rest of the paper is organized as follows. Section 2 describes the finite element modeling procedure, the near- to far-field mapping, the color-conversion scheme, as well as the randomization that suppresses diffraction effects. Section 3 describes the inverse design methodology, the optimization problem, and the numerical procedure. Section 4 discusses a range of optimized surface structures for different design goals, and Section 5 concludes on the work.

2. MODELING OF COLOR-DISPLAYING NANOSTRUCTURED SURFACES

A. Numerical Model

The numerical setup is similar to our previous study [12] but extended with a near- to far-field transformation as well as a color-conversion process. The computational domain is given in 2D and consists of three regions: an air region Ω_A; a design region Ω_D; and a bulk region Ω_B, see Fig. 1. Assuming invariance of the electromagnetic properties in the out-of-plane direction, Maxwell's equations simplify to the scalar Helmholtz equation. The scalar Helmholtz equation governs the physics for steady-state electromagnetic wave problems with a sinusoidal E_z or H_z polarized plane wave of angular frequency ω using $e^{j\omega t}$ to convert from phasor to time notation. Here we state the governing equation for an E_z polarized wave. Equivalent equations for H_z polarization are easily obtained by interchanging $E_z \leftrightarrow H_z$ and $\epsilon_r \leftrightarrow \mu_r$. The scalar Helmholtz equation is given as [20, p. 9]

$$\nabla \cdot (\mu_r^{-1} \nabla E_z) + k_0^2 \epsilon_r E_z = 0, \tag{1}$$

where ϵ_r is the relative permittivity, μ_r is the relative permeability, and $k_0 = \omega/c$ is the free-space wave number. The domain is truncated using perfectly matched layers (PML) [20] at the top and bottom boundaries. The governing equation in the Ω_{PML} regions is

$$\frac{\partial}{\partial x}\left(\frac{s_y}{s_x}\mu_r^{-1}\frac{\partial E_z}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{s_x}{s_y}\mu_r^{-1}\frac{\partial E_z}{\partial y}\right) + k_0^2 s_x s_y \epsilon_r E_z = 0, \tag{2}$$

where s_x and s_y are complex functions of the position and govern the damping properties of the PML. The incident field E_z^i is given as

$$E_z^i = E_{z0} \exp\left(-jk_0\sqrt{\mu_r\epsilon_r}\hat{\mathbf{k}}\cdot\mathbf{r}\right), \tag{3}$$

where E_{z0} is the amplitude of the wave, $\hat{\mathbf{k}} = (\hat{k}_x, \hat{k}_y)^T$ is the normalized directional wave vector, and $\mathbf{r} = (x, y)^T$ is the spatial position vector. The incident wave is generated by a surface electric current density $\mathbf{J}_s$ on Γ_i only having a z component given by (see Appendix A)

$$J_{sz} = 2\cos(\theta_{\text{in}})\sqrt{\frac{\epsilon_0\epsilon_r\mu_r}{\mu_0}}E_z^i, \tag{4}$$

where θ_{in} is the angle of the incoming wave to the normal of Γ_i. The periodic Bloch–Floquet boundary conditions on Γ_p are given by

$$E_z(x, d) = E_z(x, 0)\exp\left(-jk_0\sqrt{\mu_r\epsilon_r}d\sin(\theta_{\text{in}})\right). \tag{5}$$

The problem is solved using the finite element method (FEM) [20], and discretization details can be found in [12].

The field in close proximity and inside the nanostructure is found with the solution of the FEM problem. However, the reflected wave will most likely propagate a distance equal to many wavelengths before reaching the eye of an observer. Hence, the field of interest is the far field and not the near field found from the FEM solution. To accommodate this, a near- to far-field transformation is performed on the radial component of the scattered field based on Huygen's principle [20] and discarding any terms that decay faster than $1/(\rho)^{1/2}$ where ρ is the observation distance. The radial component of the scattered far-field E_z^{sr} can be extracted from the near field using the following expression from [20]:

$$\begin{aligned} E_z^{sr}(\rho, \lambda, \theta_{\text{out}}) \approx \sqrt{\frac{jk_0}{8\pi\rho}}\exp(-jk_0\rho)\ldots \\ \int_0^d\Bigg[(\hat{n}_y\sin(\theta_{\text{out}}) + \hat{n}_x\cos(\theta_{\text{out}}))E_z^s\ldots \\ -\frac{1}{jk_0}\left(\hat{n}_x\frac{\partial E_z^s}{\partial x} + \hat{n}_y\frac{\partial E_z^s}{\partial y}\right)\Bigg]\ldots \\ \exp(jk_0(x\cos(\theta_{\text{out}}) + y\sin(\theta_{\text{out}})))\mathrm{d}y, \end{aligned} \tag{6}$$

where $E_z^s = E_z - E_z^i$ is the scattered field and θ_{out} is the observation direction. Note that the line integral along the boundary Γ_i in Eq. (6) is only defined along the y axis because the line for Γ_i is constant with respect to x. The normalized reflection factor $L(\lambda, \theta_{\text{out}})$ in the direction θ_{out} of the scattered far-field is found as

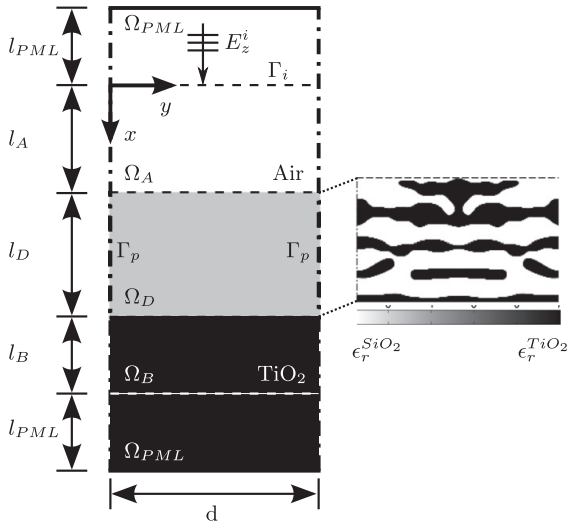


Fig. 1. Computational domain composed of an air region Ω_A, a TiO_2 region Ω_B, and a design domain Ω_D, where SiO_2 and TiO_2 is distributed to form the nanostructure. The periodicity of the structure is modeled with Bloch–Floquet boundary conditions at Γ_p.

$$L(\lambda,\theta_{\text{out}}) = \frac{\rho\lambda}{\cos(\theta_{\text{inc}})\cos(\theta_{\text{out}})}\frac{|E_z^{sr}(\rho,\lambda,\theta_{\text{out}})|^2}{|E_z^{i}(\lambda)|^2}, \tag{7}$$

where θ_{inc} is the angle of the incoming wave to the normal of the surface. In the above equation, we have multiplied with ρ in order to make the reflection factor independent of the observation distance.

The numerical model, including far-field transform, was tested against structures with analytical solutions such as a cylinder as well as the *Morpho* butterfly wing microstructure from [5].

B. Color Conversion

Since colors (of nonemitting objects) are only perceived when lit by a light source, the reflection of the dielectric nanostructure must be analyzed in a well-defined light environment. The so-called D_{65} standard illuminant spectrum that resembles daylight [21] is chosen as the light source. The reflected spectrum is then converted to coordinates in the RGB color space by applying the color-matching functions [22] shown in Fig. 2:

$$\begin{aligned} R(\theta_{\text{out}}) &= K_c\int_{380}^{760} D_{65}(\lambda)\bar{r}(\lambda)L(\lambda,\theta_{\text{out}})\mathrm{d}\lambda \\ &\approx K_c\sum_{\lambda} D_{65}(\lambda)\bar{r}(\lambda)L(\lambda,\theta_{\text{out}})\Delta\lambda, \end{aligned} \tag{8}$$

$$\begin{aligned} G(\theta_{\text{out}}) &= K_c\int_{380}^{760} D_{65}(\lambda)\bar{g}(\lambda)L(\lambda,\theta_{\text{out}})\mathrm{d}\lambda \\ &\approx K_c\sum_{\lambda} D_{65}(\lambda)\bar{g}(\lambda)L(\lambda,\theta_{\text{out}})\Delta\lambda, \end{aligned} \tag{9}$$

$$\begin{aligned} B(\theta_{\text{out}}) &= K_c\int_{380}^{760} D_{65}(\lambda)\bar{b}(\lambda)L(\lambda,\theta_{\text{out}})\mathrm{d}\lambda \\ &\approx K_c\sum_{\lambda} D_{65}(\lambda)\bar{b}(\lambda)L(\lambda,\theta_{\text{out}})\Delta\lambda, \end{aligned} \tag{10}$$

where K_c is a scaling factor and the integration is approximated by summation for 20 discrete, equidistant wavelengths between 380 and 760 nm. Note that the D_{65} spectrum is taken into account in the expression, and hence FEM simulations can be carried out with equal input intensity for each wavelength. The match among the color-matching function, reference illuminant, and light source further ensures that a reflected flat spectrum will appear white, as expected intuitively. The scaling factor K_c is found such that the RGB values from the specular reflection of a fully reflecting surface are 1. Note that this normalization does not guarantee that individual RGB numbers do not exceed unity or go below zero. This is due to the negative intervals of the color-matching functions (cf. Fig. 2).

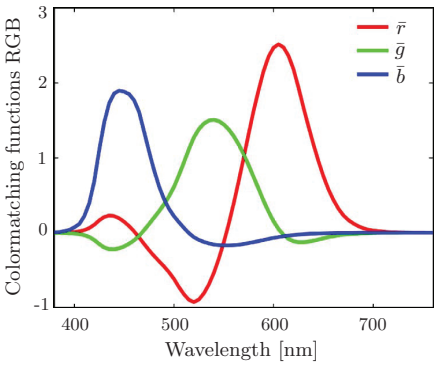


Fig. 2. Color-matching functions for finding the RGB channels of a light spectrum. The functions are derived using the 1931 CIE 2° standard observer weighting function [21] to obtain the XYZ tristimulus values and afterward converted to the RGB working space using a D_{65} reference while following ISO 22028-1:2004.

C. Suppressing Diffraction Effects

Reflection from strictly periodic structures lit by a plane wave creates diffraction. Diffraction is constructive and destructive interference, resulting in light only being reflected at certain angles. These reflection angles are wavelength-dependent. This means that a perfectly periodic structure illuminated by white light as seen in Fig. 3(a) will reflect a rainbow pattern as illustrated in the reflected spectra of Fig. 3(b). This effect can, for example, be observed when looking at a compact disc (CD). The angles of reflection are determined by the period of the surface structure. To avoid this, random height displacements can be introduced to remove the interference effects and make it possible to predict the total reflected intensity as the reflected intensity from one "period." This has been shown in theory as well as in practice in [6,23,24] and is a combination of random height variation and light incoherence, which obscure phase information in the reflected electromagnetic wave. Reference [5] demonstrates relevant random height distributions and [24] shows that a binary height distribution in practice is enough if the wanted color can be represented by a small band of wavelengths. The random height interval can in general be kept below half of the longest

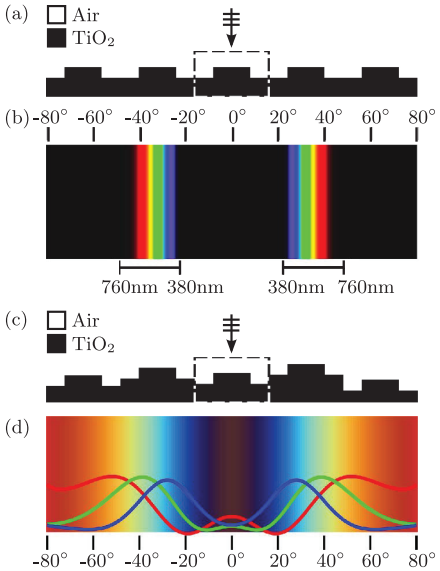


Fig. 3. Perfectly periodic nanostructure (a) with corresponding diffraction pattern (b) and a height randomized structure (c) with the diffraction effect eliminated (d). The display colors are exaggerated for the purpose of illustration. The area surrounded by the dashed line corresponds to one "unit cell."

wavelength of interest corresponding to a phase change of 360° at that wavelength.

The simulations are carried out assuming that a random variation can be applied afterward (see, e.g., [23]) as illustrated in Fig. 3(c), giving rise to the reflected spectra seen in Fig. 3(d), where all diffraction effects are removed. We will not deal with this aspect in further detail here, but assume that a random height variation as seen in [5] can be applied to make the response from one period of the structure appear similar to the response of an array of randomized cells.

3. INVERSE DESIGN METHODOLOGY

A. Material Distribution by Topology Optimization

Light can be reflected and/or refracted at interfaces between two media, depending on the spatial placement and distribution of the material properties (ϵ_r and μ_r). Hence, control of these material properties allows wave manipulation for various purposes, e.g., creating reflection bands for coloration. Here we work with a TiO_2 substrate with a nanostructure composed of TiO_2 and SiO_2; however, the developed topology optimization methodology can handle any simple (linear, homogeneous, isotropic) material combinations including metallic structures [25] and also frequency-dependent material properties. The method works by varying the distribution of materials within a bounded design domain in order to optimize certain responses of the physical system. In the design domain Ω_D between the air and TiO_2 substrate, the relative permittivity can be varied continuously on an element basis between the permittivity for SiO_2 ($\epsilon_r^{SiO_2} = 2.25$) and TiO_2 ($\epsilon_r^{TiO_2} = 7.02$). We employ a standard density-based topology optimization method [16] and restrict our investigations to nonmagnetic materials ($\mu_r = 1$). A continuous design variable $\gamma \in [0;1]$ is introduced for each element in the design domain Ω_D and controls the element material properties. Here $\gamma_e = 0$ corresponds to SiO_2 and $\gamma_e = 1$ corresponds to TiO_2:

$$\epsilon_r(\gamma_e) = \epsilon_r^{SiO_2} + \gamma_e(\epsilon_r^{TiO_2} - \epsilon_r^{SiO_2}). \tag{11}$$

The continuous design variable formulation allows us to solve the optimization problem with efficient gradient-based design updates [26]. In principle, the optimization may result in "gray scale" results, i.e., elements that neither correspond to SiO_2 nor to TiO_2; however, the robust design formulation developed in [27,28] and also used in [12] ensures almost discrete designs through a continuation strategy. Symmetric solutions are ensured by mirroring of design variables.

B. Optimization Problem

The idea of the inverse design methodology is to find a nanostructure that displays a desired color effect in a specific angle or angular range. The desired color is given as a reference in the RGB color space $RGB_r = [R_r, G_r, B_r]$. The direction of the color vector $RGB(\theta_{out})$ (for nonzero length) decides the color, and the length decides the color intensity. In order to create a desired color effect, it is crucial that the color vector from the nanostructure is parallel to the reference color vector. This is ensured by imposing a constraint on the normalized cross-product between the actual and the prescribed color vectors. With the color vector constrained in the desired direction in the color space, the normalized intensity of the reflected light is maximized. Furthermore, a volume constraint is imposed to prevent congestion in the design domain. The optimization problem is formulated as a maxmin formulation for design problems involving a color effect in an angular range or different colors reflected in different directions simultaneously. The optimization maxmin problem maximizing the minimum intensity of several angular directions is formulated as

$$\begin{aligned} \max_{\gamma} \quad & \Phi := \min_{k=1,\ldots,N} \frac{|RGB(\theta_{\text{out}}^k)|^2}{|RGB_r^k|^2}, \\ \text{s.t.} \quad & \frac{|RGB(\theta_{\text{out}}^k) \times RGB_r^k|^2}{|RGB_r^k|^2} \le \tau^2, \qquad k = 1,\ldots,N, \\ & \frac{1}{V_{\Omega_D}} \int_{\Omega_D} \gamma \mathrm{d}\Omega_D - \beta \le 0, \\ & 0 \le \gamma \le 1, \end{aligned} \tag{12}$$

where k is the index of the N angular directions, τ is the highest admissible value of the error between the actual and prescribed color vectors, V_{Ω_D} is the total volume of the design domain Ω_D, and β is the admissible volume fraction. The constraint on the color vector is introduced for each angular direction in the problem. The prescribed reference colors in the color constraints do not need to be the same but can be set to any desired RGB value. The average of the RGB values obtained from both E_z and H_z polarized waves is used in Eq. (12) for unpolarized color effects.

The design is updated iteratively using the gradient-based optimization routine method of moving asymptotes (MMA) [29]. The normalized intensity given as Φ in Eq. (12) is used as a performance measure for the nanostructures, and the sensitivities are obtained using the adjoint method [15]. The maxmin problem in Eq. (12) is in general a challenging problem due to non-differentiability of the min-function. To circumvent this, the problem is reformulated to the so-called bound-formulation, which is a standard conversion allowed by the MMA optimizer. The interested reader is referred to the MMA paper [29] and Svanberg's publicly available codes for further explanations.

The design procedure for black is constructed as a special case. The objective for black is to absorb or transmit the light for all visible wavelengths and thereby have zero reflection. Hence the maximum intensity $\Phi(\theta_{\text{out}})$ at several angular directions is minimized in the optimization problem [cf. Eq. (12)] while using white as a reference color ($R_r = 1$, $G_r = 1$, $B_r = 1$). The lower the value of Φ, the better the nanostructure is in displaying black.

The proposed procedure is implemented in MATLAB utilizing parallel computations for the frequency sweeps. Convergence is typically reached after 500–1000 material redistribution steps, where the main computational effort in each step consists in the forward and adjoint FEM solutions for the 20 frequencies. This effort may be significantly reduced using Padé approximants for the frequency sweeps [15], a feature we will implement in future versions of the code. Presently, one optimization runs overnight on an eight-processor computer.

4. STRUCTURAL COLORS GENERATED BY TOPOLOGY OPTIMIZED NANOSTRUCTURES

In this section, we demonstrate the procedure's ability to design nanostructured surfaces with tailored color properties for narrow and wider viewing angles, for polarized and

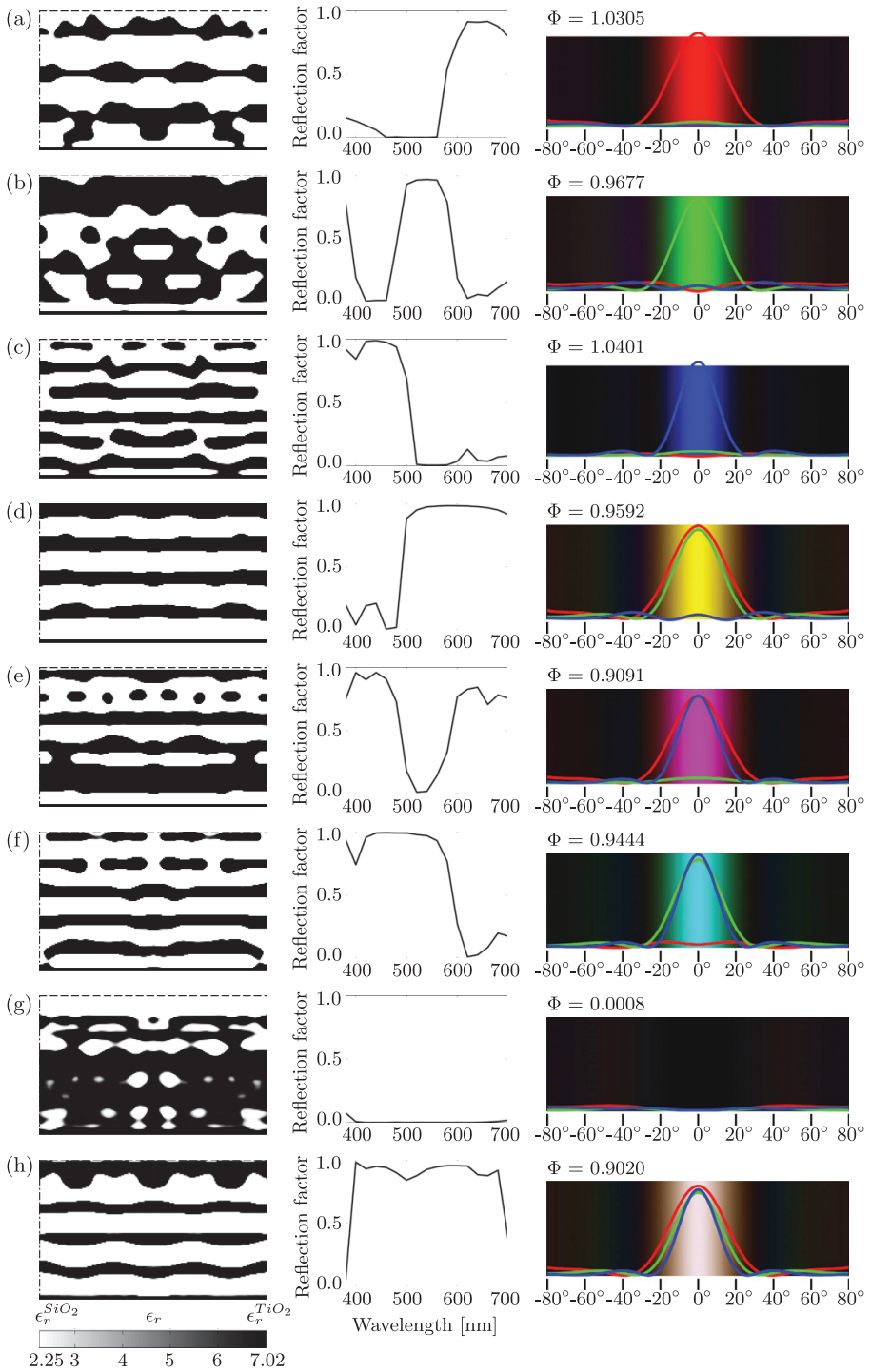


Fig. 4. Color optimization for 0° observation angle. The left column shows the material distribution; the center column the frequency response seen from the observer; and the right column shows the angular color spectrum including individual color distributions. Structures (a)–(h) are optimized for red ($RGB_r = [1, 0, 0]$), green [0,1,0], blue [0,0,1], yellow [0,1,1], magenta [1,0,1], cyan [1,1,0], black [0,0,0], and white [1,1,1], respectively.

non-polarized light, as well as for prescribed iridescent, i.e., angular dependent, color properties. The nanostructuring is obtained by a nanostructured mix of SiO_2 and TiO_2 on a TiO_2 substrate.

We consider a periodically repeated symmetric design domain as indicated with gray in Fig. 1. The period is $d = 1000$ nm and the thickness is $l_D = 600$ nm. The design domain is discretized with 120×200 finite elements and a maximum of

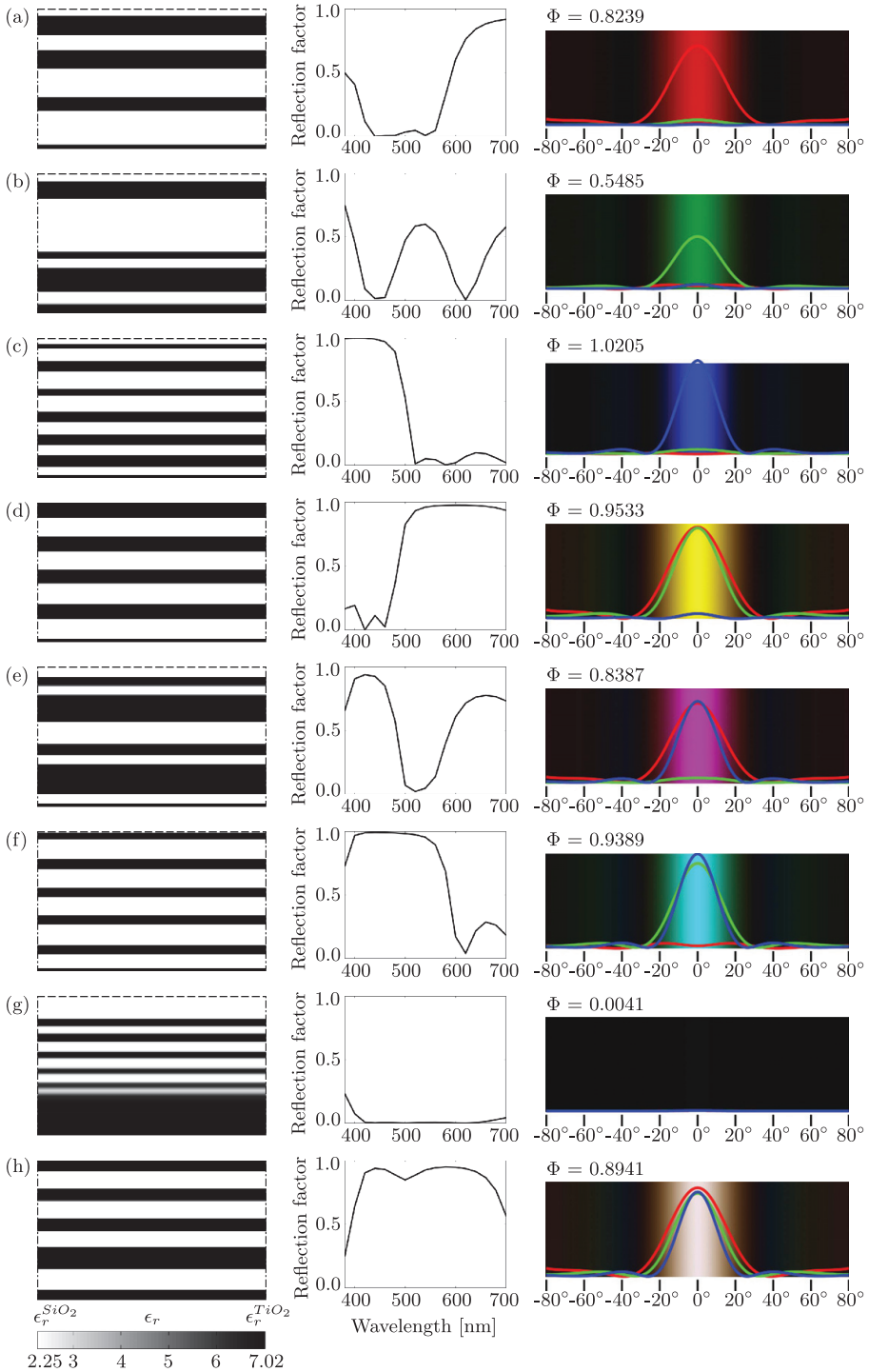


Fig. 5. Color optimization for 0° observation angle but design freedom limited to simple layered Bragg-like structures. Other details can be read from the caption of Fig. 4.

$\beta = 65\%$ of TiO_2 ($\epsilon_r^{TiO_2} = 7.02$) can be distributed in the SiO_2 ($\epsilon_r^{SiO_2} = 2.25$) base material. The choice of design domain thickness l_D is chosen arbitrarily. A very small thickness will not be able to produce interesting color effects whereas a large thickness (i.e., much larger than the wavelength) will result in intense colors and be of less challenge to the optimizer.

We have used $\tau = 0.05$ as default value in Eq. (12), i.e., we allow a 5% deviation from the desired color vector. Choosing a

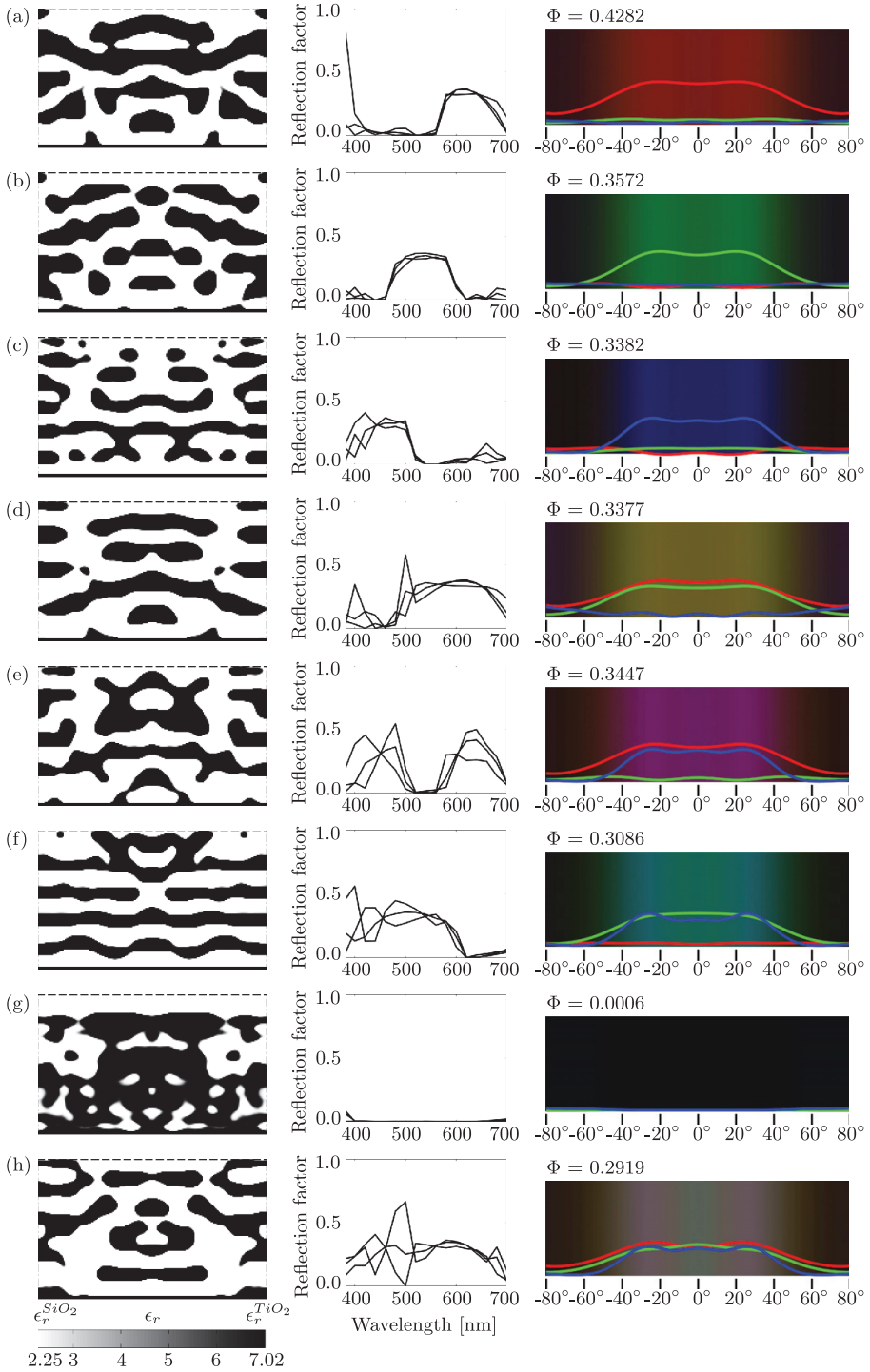


Fig. 6. Color optimization for –30° to 30° observation angle with full design freedom. Other details can be read from the caption of Fig. 4. The three curves in the reflection plots correspond to the reflection spectra for the three optimization angles (0°, 15°, and 30°).

Table 1. Objective Values for RGB Optimizations Corresponding to 0°, 0° Constrained to Simple Layerings and Wider Angular Spectrum [−30°, 30°][a]

Objective	Figure	R	G	B	Y	M	C	W	B
0°	4	1.03	0.97	1.04	0.96	0.91	0.94	0.90	0.00
0° layer	5	0.82	0.55	1.02	0.95	0.83	0.93	0.89	0.00
[−30°, 30°]	6	0.43	0.36	0.34	0.34	0.34	0.31	0.29	0.00

[a]The objective values correspond to Figs. 4, 5, and 6, respectively. Colors are abbreviated red (R), green (G), blue (B), yellow (Y), magenta (M), cyan (C), black (B), and white (W).

smaller value will make it harder for the optimizer to converge and will most probably not be visually perceptible anyway. As initial designs, we have used both pure SiO_2 as well as a Bragg grating designed to have a reflection band at a wavelength corresponding to the desired color. When running the optimization for different initial designs we minimize the risk of finding a less desirable local optimum. In each case, we show the optimized design with the best performance out of the two initial designs. In most cases, the difference in performance is small anyway.

To begin with, we consider three different cases all with normal (0°) incident light: (1) optimizing the color in the

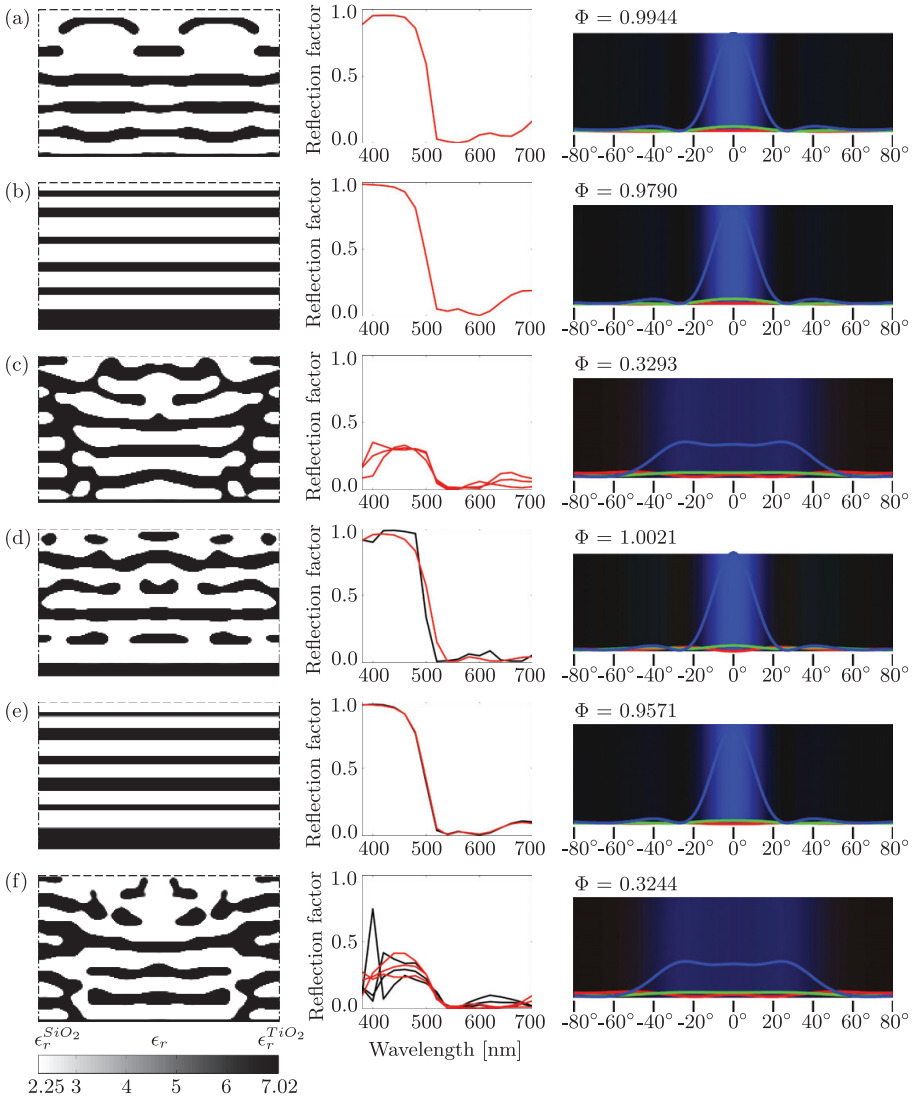


Fig. 7. Color optimization for combined E_z and H_z polarization. (a)–(c) Structures optimized for H_z polarization corresponding to those optimized for E_z in Figs. 4(c), 5(c), and 6(c), respectively. (d)–(f) Similar structures optimized for combined E_z and H_z polarization. Red curves in the reflection plots correspond to H_z polarization and black to E_z.

specular direction (0°) with free topology; (2) optimizing the color in the specular direction (0°) with the geometry restricted to simple (Bragg grating-like) layerings; and (3) optimizing the color in the angular range (−30° to +30°) with free topology. In practise, the optimizations for angular intervals are only optimized for N discrete angles (here $N = 5$, which is reduced to $N = 3$ for symmetry reasons). For the three formulations, we optimize structured surfaces to produce the six main colors: red, green, blue, magenta, yellow, and cyan as well as black and white. The results are shown in Figs. 4, 5, and 6, respectively. The result of each individual optimization is presented as a bitmap picture of the optimized topology (left panel), reflection spectra for each output angle and polarization (middle panel), and an angular color spectrum with background color corresponding to the perceived color and individual curves for the three basis colors: red, green, and blue (right panel). The Φ values given for each example denotes the color intensity as defined in Eq. (12).

The optimization problem is highly nonconvex, and convergence to a global optimum cannot be guarantied. However, a continuation strategy runs with different starting guesses, and previous experience make us confident that very good solutions are found. In order to perform a sanity check of the obtained results, we compare objective functions for the three cases in Table 1. As expected, the reduced design freedom, as represented by the layered structures in Fig. 5, cannot in all cases produce the same intensity as the freely optimized structures from Fig. 4. Likewise, the structures optimized for a wider angular interval (Fig. 6) show much lower angular color intensity. It also can be seen that the color intensities for green and magenta are lower than for the others colors, especially for the simple layered structures. This also is to be expected since the forming of green and magenta requires a stop band and a pass band, respectively, whereas all other colors can be obtained by simpler low or high-pass filters.

From the comparisons above, it is clear that freely optimized structures in Fig. 4 are indeed better than simple layered ones in Fig. 5. With the increased design freedom, it is possible not only to create simple 1D gratings with specific pass and stop behavior but also to scatter light away from the observation point or alternatively to create more complex structures that may improve transmission of frequencies unwanted in the reflection spectrum. Both possibilities may lead to improved color response at the observation point.

The red, green, and magenta designs perform surprisingly bad for the simple layered designs in Fig. 5. However, these are also the designs that resemble layered structures the least for the full design freedom in Fig. 4. Hence, it is concluded that complex structures seen in Figs. 4(a), 4(b), and 4(e) are indeed needed to ensure intense and pure colors.

For the structures optimized for broad angular response in Fig. 6, the simple layered structures become highly suboptimal since they only reflect light back in a limited angular range. Instead all the optimized designs exhibit rather complex structures that are difficult to interpret. They all exhibit a kind of directional photonic crystal structure, and it is observed that the periodicity varies with wavelength. For example, the structure optimized for red color generation (long wavelength) has larger and more distant inclusions than the structure optimized for blue color generation (short wavelength). For the chosen cell size, the angle for first-order diffraction for red light is 43°, and for blue light it is 27°. Hence, when optimizing color for the angular interval [−30°, 30°], red will have an advantage since it does not need to include second-order diffraction effects in order to reflect evenly in a broad angular interval. This observation may explain the superiority of the red design in Fig. 6 compared to the other colors.

The previous designs were all obtained for a single polarization. In Fig. 7 we demonstrate the effect of polarization

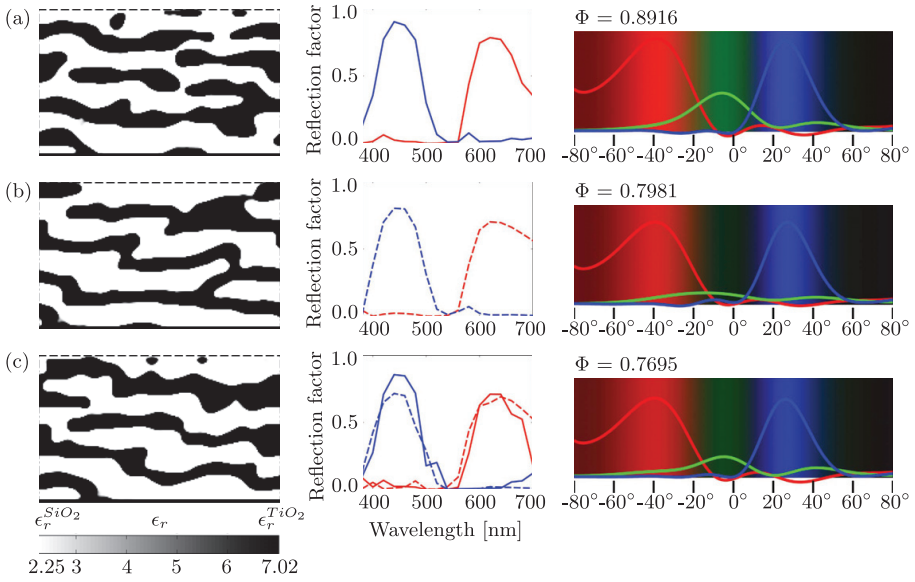


Fig. 8. Optimization for prescribed iridescent colors. Red color [1,0,0] intensity is optimized for −43° observer direction, and blue color intensity [0,0,1] is optimized for 27° observer direction. Designs (a)–(c) correspond to optimized structures for E_z, H_z, and combined E_z to H_z polarization, respectively. Red curves correspond to the reflection in the −43° direction and blue to the reflection in the 27° direction.

on structures optimized for blue color generation. Figures 7(a)–7(c) show the designs optimized for H_z polarization, which can be compared to the similar ones optimized for E_z polarization in Figs. 4(c), 5(c), and 6(c), respectively. These designs appear not to be qualitatively much different from each other, indicating that polarization is not a big issue from a design view point. Figures 7(d)–7(f) show the corresponding designs optimized for both polarizations simultaneously. Color intensities only drop insignificantly, indicating that it is quite easy to design polarization-independent structural color surfaces using the developed methodology.

The final example considers the design of a surface with angular-dependent (iridescent) color properties. Again, the incoming light is perpendicular to the surface, but we require the reflected colors to be red in the −43° direction and blue in the 27° direction. The optimized structures for E_z and H_z polarization are shown in Figs. 8(a) and 8(b), respectively. It is observed that large red and blue color intensities can be obtained in the prescribed directions, and that the color is vague green in between. If the green color is unwanted, one may prescribe black for the zero-degree reflection angle, making the optimization generate a structure that transmits more of the green color instead of reflecting it. Figure 8(c) shows the same optimization performed for simultaneous E_z and H_z polarization. Again, requirements for nonpolarized response do not alter the obtainable color intensities significantly. The optimized structures are difficult to interpret; however, slanted layered structures can be seen that help to redirect the waves in the desired directions.

5. CONCLUSION

We have proposed a systematic procedure for generating nanostructured surfaces with prescribed structural color response. The procedure is demonstrated for a number of cases, including generation of arbitrary colors in narrow- or wide-angle intervals as well as for angular-dependent or iridescent color generation. The developed software is general and can be applied to any kind of material properties. For the examples, we considered a TiO_2/SiO_2 nanostructure on a TiO_2 substrate, but the procedure also can handle single material structures in air. In this case, manufacturing constraint must be included to prevent nonconnected parts as, for example, seen in [12]. In the present paper, we allow full design freedom, which results in strong color effects even for thin design domains. If complex nanostructures are not allowed by the manufacturing process, additional constraints must be imposed, inevitably resulting in decreased color-generation performance. The proposed procedure can find many applications. It may be used for the optimal study of structural color surfaces found in nature, extended to three dimensions, allowing for more advanced color effects, and extended to plasmonic design problems for color generation (cf. [30]) or thin film solar cells [31,32].

APPENDIX A: DERIVATION OF IMPRESSED SURFACE ELECTRIC CURRENT DENSITY

On the incoming wave boundary Γ_i, we have the boundary condition [20, p. 10]

$$\hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{J} = \hat{\mathbf{z}} J_{sz}, \tag{A1}$$

where $\hat{\mathbf{n}} = \hat{\mathbf{x}}\hat{n}_x + \hat{\mathbf{y}}\hat{n}_y$ is the downward pointing normal to Γ_i, and $\mathbf{H}_2/\mathbf{H}_1$ are the magnetic fields existing just above/below Γ_i, respectively, and J_{sz} is the impressed surface electric current density, which only has a z component since the H fields and the normal are perpendicular to $\hat{\mathbf{z}}$. For simple media, it holds that $\mathbf{H} = 1/(j\omega\mu_r\mu_0)\nabla \times \mathbf{E}$, and since we have defined $\mathbf{E} = \hat{\mathbf{z}}E_z$, this in turn leads to

$$\nabla \times \mathbf{E} = \hat{\mathbf{x}}\frac{\partial E_z}{\partial y} - \hat{\mathbf{y}}\frac{\partial E_z}{\partial x}. \tag{A2}$$

The cross-product $\hat{\mathbf{n}} \times \mathbf{H}$ from this can be found to be

$$\hat{\mathbf{n}} \times \mathbf{H} = \frac{1}{j\omega\mu_r\mu_0}\left(\hat{n}_y\frac{\partial E_z}{\partial y} + \hat{n}_x\frac{\partial E_z}{\partial y}\right) = \frac{1}{j\omega\mu_r\mu_0}\hat{\mathbf{n}} \cdot (\nabla E_z), \tag{A3}$$

where μ_r is assumed constant around Γ_i. Combining Eqs. (A1) and (A3), the expression

$$\hat{\mathbf{n}} \cdot (\nabla E_{z2} - \nabla E_{z1}) = \hat{\mathbf{z}} j\omega\mu_r\mu_0 J_{sz} \tag{A4}$$

is obtained, where E_{z1}, E_{z2} are the z components of $\mathbf{E}_1$, $\mathbf{E}_2$ related to $\mathbf{H}_1$, $\mathbf{H}_2$, respectively. Since we cannot expect to excite a wave in one direction without exciting one in the opposite direction as well, we define the fields that should be excited by J_{sz} to be

$$\begin{aligned}\mathbf{E}_1 &= \hat{\mathbf{z}}E_{z0}\exp\left(-jk_0\sqrt{\mu_r\epsilon_r}\hat{\mathbf{k}}_1 \cdot \mathbf{r}\right),\\ \mathbf{E}_2 &= \hat{\mathbf{z}}E_{z0}\exp\left(-jk_0\sqrt{\mu_r\epsilon_r}\hat{\mathbf{k}}_2 \cdot \mathbf{r} + j\phi\right),\end{aligned} \tag{A5}$$

where ϕ is a constant chosen such that $\mathbf{E}_1$, $\mathbf{E}_2$ match on Γ_i. Assuming that Γ_i is parallel to the y direction and defining $\hat{\mathbf{k}}_1 \equiv \hat{\mathbf{k}} = (\hat{k}_x, \hat{k}_y)^T$ such that the incoming wave is equivalent to Eq. (3), then $\hat{\mathbf{k}}_2 = (-\hat{k}_x, \hat{k}_y)^T$. This in turn means that $\nabla E_{z1} = -jk_0\sqrt{\mu_r\epsilon_r}\hat{\mathbf{k}}E_{z1}$ and $\nabla E_{z2} = -jk_0\sqrt{\mu_r\epsilon_r}\hat{\mathbf{k}}_2 E_{z2}$, which, combined with Eq. (A4), leads to

$$J_{sz} = \frac{jk_0\sqrt{\mu_r\epsilon_r}}{j\omega\mu_0\mu_r}\hat{\mathbf{n}} \cdot (\hat{\mathbf{k}}_1 E_{z1} - \hat{\mathbf{k}}_2 E_{z2}). \tag{A6}$$

Since the two electrical fields are the same on Γ_i and $\hat{\mathbf{n}} = (1, 0)$ for Γ_i along the y axis, the surface electric current density exciting the wave in Eq. (3) (and also a wave in the other direction) can then be found as

$$J_{sz} = 2\hat{k}_x\sqrt{\frac{\epsilon_0\epsilon_r\mu_r}{\mu_0}}E_{z,1}. \tag{A7}$$

ACKNOWLEDGMENTS

We are grateful for the support from the Danish National Technology Foundation through the projects NANOPlast and ODAAS and the Villum Foundation through the NextTop project.

REFERENCES

1. M. Kolle, *Photonic Structures Inspired by Nature* (Springer, 2011).

2. S. Kinoshita and S. Yoshioka, "Structural colors in nature: the role of regularity and irregularity in the structure," Chem. Phys. Chem. **6**, 1442–1459 (2005).
3. M. Crne, V. Sharma, J. Blair, J. O. Park, C. J. Summers, and M. Srinivasarao, "Biomimicry of optical microstructures of Papilio palinurus," Europhys. Lett. **93**, 14001 (2011).
4. S. Kinoshita, S. Yoshioka, and J. Miyazaki, "Physics of structural colors," Rep. Prog. Phys. **71**, 076401 (2008).
5. R. T. Lee and G. S. Smith, "Detailed electromagnetic simulation for the structural color of butterfly wings," Appl. Opt. **48**, 4177–4190 (2009).
6. M. A. Steindorfer, V. Schmidt, M. Belegratis, B. Stadlober, and J. R. Krenn, "Detailed simulation of structural color generation inspired by the Morpho butterfly," Opt. Express **20**, 21485–21494 (2012).
7. N. Okada, D. Zhu, D. Cai, J. B. Cole, M. Kambe, and S. Kinoshita, "Rendering Morpho butterflies based on high accuracy nano-optical simulation," J. Opt. **42**, 25–36 (2013).
8. A. Saito, "Material design and structural color inspired by biomimetic approach," Sci. Technol. Adv. Mat. **12**, 064709 (2011).
9. K. Kumar, H. Duan, R. S. Hegde, S. C. W. Koh, J. N. Wei, and J. K. W. Yang, "Printing colour at the optical diffraction limit," Nat. Nanotechnol. **7**, 557–561 (2012).
10. D. C. Dobson, "Optimal design of periodic antireflective structures for the Helmholtz equation," Euro. J. Appl. Math. **4**, 321–339 (1993).
11. K. Fuchi, A. R. Diaz, E. Rothwell, R. Ouedraogo, and A. Temme, "Topology optimization of periodic layouts of dielectric materials," Struct. Multidiscip. Optim. **42**, 483–493 (2010).
12. K. S. Friis and O. Sigmund, "Robust topology design of periodic grating surfaces," J. Opt. Soc. Am. B **29**, 2935–2943 (2012).
13. M. P. Bendsøe and O. Sigmund, *Topology Optimization—Theory, Methods and Applications* (Springer-Verlag, 2004).
14. M. P. Bendsøe and N. Kikuchi, "Generating optimal topologies in structural design using a homogenization method," Comput. Methods Appl. Mech. Eng. **71**, 197–224 (1988).
15. J. S. Jensen and O. Sigmund, "Topology optimization of photonic crystal structures: a high-bandwidth low-loss T-junction waveguide," J. Opt. Soc. Am. B **22**, 1191–1198 (2005).
16. J. S. Jensen and O. Sigmund, "Topology optimization for nano-photonics," Laser Photon. Rev. **5**, 308–321 (2011).
17. Y. Elesin, B. S. Lazarov, J. S. Jensen, and O. Sigmund, "Time domain topology optimization of 3D nanophotonic devices," Photon. Nanostruct. Fundam. Appl., doi: 10.1016/j.photonics.2013.07.008 (2013).
18. A. Erentok and O. Sigmund, "Topology optimization of sub-wavelength antennas," IEEE Trans. Anten. Propag. **59**, 58–69 (2011).
19. A. R. Diaz and O. Sigmund, "A topology optimization method for design of negative permeability metamaterials," Struct. Multidiscip. Optim. **41**, 163–177 (2010).
20. J. Jin, *The Finite Element Method in Electromagnetics* 2nd ed. (Wiley, 2002).
21. CIE, "Selected colorimetric tables," http://www.cie.co.at.
22. R. S. Berns, F. W. Billmeyer, and M. Saltzman, *Billmeyer and Saltzman's Principles of Color Technology* (Wiley-Interscience, 2000).
23. A. Saito, Y. Miyamura, Y. Ishikawa, J. Murase, M. Akai-Kasaya, and Y. Kuwahara, "Reproduction, mass production, and control of the Morpho butterfly's blue," in *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics II*, T. J. Suleski, W. V. Schoenfeld, and J. J. Wang, eds. (SPIE, 2009), Vol. **7205**, p. 720506.
24. A. Saito, M. Yonezawa, J. Murase, S. Juodkazis, V. Mizeikis, M. Akai-Kasaya, and Y. Kuwahara, "Numerical analysis on the optical role of nano-randomness on the Morpho butterfly's scale," J. Nanosci. Nanotechnol. **11**, 2785–2792 (2011).
25. J. A. Andkjær, S. Nishiwaki, T. Nomura, and O. Sigmund, "Topology optimization of grating couplers for the efficient excitation of surface plasmons," J. Opt. Soc. Am. B **27**, 1828–1832 (2010).
26. O. Sigmund, "On the usefulness of non-gradient approaches in topology optimization," Struct. Multidisc. Optim. **43**, 589–596 (2011).
27. O. Sigmund, "Manufacturing tolerant topology optimization," Acta Mech. Sinica **25**, 227–239 (2009).
28. F. Wang, B. S. Lazarov, and O. Sigmund, "On projection methods, convergence and robust formulations in topology optimization," Struct. Multidisc. Optim. **43**, 767–784 (2011).
29. K. Svanberg, "The method of moving asymptotes—a new method for structural optimization," Intern. J. Numer. Meth. Eng. **24**, 359–373 (1987).
30. K. Kumar, H. Duan, R. S. Hegde, S. C. W. Koh, J. N. Wei, and J. K. W. Yang, "Printing colour at the optical diffraction limit," Nat. Nanotechnol. **7**, 557–561 (2012).
31. M. B. Dühring and O. Sigmund, "Optimization of extraordinary optical absorption in plasmonic and dielectric structures," J. Opt. Soc. Am. B **30**, 1154–1160 (2013).
32. X. Sheng, S. G. Johnson, J. Michel, and L. C. Kimerling, "Optimization-based design of surface textures for thin-film Si solar cells," Opt. Express **19**, A841–A850 (2011).

Publication P3

Optical role of randomness for structured surfaces

Optical role of randomness for structured surfaces

Villads Egede Johansen

Technical University of Denmark, Department of Mechanical Engineering, 2800 Kgs. Lyngby, Denmark (vejo@mek.dtu.dk)

Received 2 January 2014; revised 2 January 2014; accepted 26 February 2014;
posted 5 March 2014 (Doc. ID 198272); published 7 April 2014

It has long been known that random height variations of a repeated nanoscale structure can give rise to smooth angular color variations instead of the well-known diffraction pattern experienced if no randomization is present. However, until now there have been few publications trying to explain this and similar phenomena taking outset in electromagnetic theory. This paper presents a method for analyzing far-field reflection from a surface constructed by translated instances of a given structure. Several examples of the effect of random translations are given. © 2014 Optical Society of America

OCIS codes: (030.1670) Coherent optical effects; (050.1960) Diffraction theory; (240.6700) Surfaces; (290.0290) Scattering; (330.1690) Color; (330.7326) Visual optics, modeling.
http://dx.doi.org/10.1364/AO.53.002405

1. Introduction

In the 17th century, Robert Hooke discovered how dielectric structures with size features comparable to the wavelengths of light were an important part of the color appearance of certain animals he studied under a microscope [1]. However, the theory of light was not well developed at the time; there were no computers, and the optical microscopes simple, so the field of structural colors remained rather untouched. In recent decades, the invention of electron microscopes, the computer, and the wave theory of light in the last century have made it possible to do more in-depth investigation of this field that has more than 500 million years of history in nature [2].

Controlling light reflection by interaction with structures is crucial for many applications. Retroreflectors, aluminum foil, solar cells, and security holograms are just a few examples encountered frequently in everyday life. These examples also show that there are many motivations for controlling color appearance of an object besides the visual appearance, and that improving the understanding of light's interaction with surfaces can improve a wide range of engineering applications and possibly initiate new inventions.

1559-128X/14/112405-11$15.00/0

Many important contributions to the understanding of structural colors have been uncovered by studying the nanostructure of the wing of the Morpho rhetenor butterfly. The results range from what is presented in [3] and up until present day, where it is possible to make a 3D rendered model of the wing's color appearance and reflection based on modeling of the structure of the wing [4]. On the Morpho type in general, excellent works have been published dating longer back, see, e.g., [5,6]. The reasons for why this structure especially has become so central in the analysis of structural colors are probably (1) the fact that its structure is more or less invariant along one axis, leaving it possible to simulate only a cross section of the model; (2) the rather simple shape of the structure, making it possible at an early stage to obtain good results just by analyzing it as a multilayer structure and then extend the model more and more to refine the results.

One of the properties that still needs investigation is how to model random displacements of the individual, repetitive structures present in the Morpho rhetenor's wing (see Fig. 3.45 in [7] and the following description in [8]). By numerical as well as practical experiments, it has been shown how these random displacements of a periodic structure seem to smooth out the otherwise strong diffractive effects, which are expected from reflections of periodic structures. A method for incorporating randomness without doing

simulations using a large numerical domain is still missing. In this paper, we will present a method for analyzing repeated structures with (or without) random translations of the elements. As a benchmark, previously published results on the random behavior related to the Morpho rhetenor butterfly will be used, but the method has a much broader aim than this: it should make it possible to design structures with new color effects, and it should help in giving a better intuitive understanding of the influence of different kinds of randomization of structures (e.g., in-plane versus height displacements). This paper focuses on visible light, but the method is applicable in all parts of the electromagnetic spectrum.

The rest of the paper is organized as follows. Section 2 motivates the work in this paper by showing examples in the literature where it can be used. Section 3 presents the background needed to analyze the random effects. Section 4 applies the knowledge to results obtained in earlier studies and shows how these could have been predicted using the presented method. Section 5 gives some examples on what influence the randomness would have on the color appearance of a surface for some specific cases, and finally Section 6 concludes on the presented results.

2. Observations of Randomness in the Literature

In this paper, we will focus on surfaces comprised of repeated unit structures in the x, y plane with some per unit height displacement in the z direction (see Fig. 1 for an example of a repeated structure). Structures fitting to this description are found in many places in the literature of optics. Partly in cases where the translations of the copied structures are deterministic and partly where the description contains random parameters. To motivate the present paper, some publications relying on one unit structure repeated throughout a domain are listed below:

- Several publications exist considering the influence of randomness of binary gratings: in [9] an experimental study of in-plane randomness is conducted with results that can also be explained from the present paper; and in [10] some binary height variations of gratings are studied that also can be analyzed using the present paper.
- Phase gratings using the properties of randomness for concrete products such as, e.g., a surface giving *controlled angular redirection of light* for windows to improve indoor lighting environment has been presented in [11]. Here the effect of in-plane randomness is observed—but not explained.
- Designing random disorder is also seen in photovoltaic solar cell applications as presented in, e.g., [12]. In [12], the method presented in the present paper could have been used to investigate the effect of different randomization parameters before utilizing full wave simulations for a detailed study.

For testing the usefulness of the presented method, studies of the nanostructure of the Morpho butterfly's wing will be used, since it has undergone many studies in the last decades with focus on different aspects of randomness. The works to be used are:

- "Detailed electromagnetic simulation for the structural color of butterfly wings [13]," which shows numerically how the far-field response of one lit Morpho butterfly ridge (which is the "unit structure" of this butterfly's wing) almost corresponds to the response of many random height translated elemenents;
- "Reproduction, mass-production, and control of the Morpho-butterfly's blue [14]," which shows by experiment how a binary random pattern with a structure on top can be used to generate a smooth color effect compared to no randomization;
- "Numerical analysis on the optical role of nano-randomness on the Morpho butterfly's scale [15]," which conducts several numerical experiments with different kinds of randomization parameters to investigate the effect of these, and;
- "Detailed simulation of structural color generation inspired by the Morpho butterfly [16]," which—among other numerical experiments—contains an analysis of the effect of different maximum heights chosen for randomization.

The method to be presented will be confirmed by these observations.

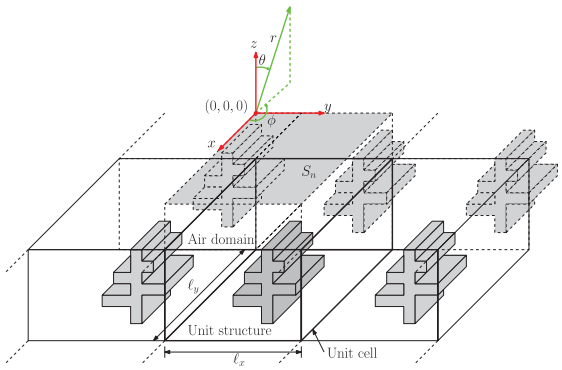


Fig. 1. Same structure divided into cells with an air domain above.

3. Method for Analyzing Random Translations

In this section, a method for analyzing (random) translations of structures is presented. It is described for a 3D general case, even though the examples later on will be 2D. The analysis is performed for time harmonic waves using the time factor $e^{j\omega t}$.

A. Huygens' Principle

Consider an electromagnetic structure with its volumes divided into different cells—which in this article will be referred to as *units* when all cells contain the same structure—as seen in Fig. 1. By Huygens' principle and the image principle [17], it is possible to calculate the (near-field as well as) far-field contribution in the upper hemisphere from

any structure by considering the so-called equivalent surface currents, $\mathbf{J}_e$, calculated on an infinite plane above the structure. An example of this could be the structure seen in Fig. 1 if the plane S_n is extended to infinity. Due to the linearity of Maxwell's equations, the plane can be split into parts with each of their far-field contribution, and by summing them we still arrive at the same total far-field contribution.

Now by defining a plane above the structure in Fig. 1 and splitting it into parts such that each part follows the projection of the cell on the plane—like S_n in Fig. 1—the magnetic far-field contribution for the **H**-field, $\mathbf{H}^f$, can be found as (see, e.g., Chapter 6 in [17])

$$\mathbf{H}^f(\theta,\phi) = \sum_{n\in\mathbb{N}} \mathbf{H}_n^f(\theta,\phi), \tag{1}$$

where

$$\mathbf{H}_n^f(\theta,\phi) = -jk\frac{e^{-jkr}}{4\pi r}\hat{\mathbf{r}} \times \int_{S_n} \mathbf{J}_e e^{jk\hat{\mathbf{r}}(\theta,\phi)\cdot\mathbf{r}_o} \mathrm{d}S_n, \tag{2}$$

where $k = 2\pi/\lambda$ is the wavenumber, the S_n's ($n \in \mathbb{N}$) are all unique parts making up the total surface S, r is the distance from an arbitrarily located origin on S_n to the evaluation point (since r is used in places where phase information is not important, it is assumed constant), $\hat{\mathbf{r}} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the direction toward the far-field evaluation point, and $\mathbf{r}_0 = (x,y,z) \in S_n$ is the position vector (as measured from the origin) to a point on S_n. The variables x,y,z,θ,ϕ are all defined as in Fig. 1.

In the far-field and free space, the radiated wave locally approaches a TEM (transverse electromagnetic) wave, and one can therefore make use of the relation between the electric and magnetic field due to this behavior:

$$\mathbf{E}^f = \eta\mathbf{H}^f \times \hat{\mathbf{r}}, \tag{3}$$

where $\eta \approx 377\Omega$ is the free-space impedance, to calculate the irradiance E—which equals the magnitude of Poynting's vector—as

$$E = \frac{1}{2}\left|\mathbf{E}^f \times \overline{\mathbf{H}^f}\right| = \frac{1}{2}\eta|\mathbf{H}^f|^2 = \frac{1}{2}\eta\left|\sum\nolimits_{n\in\mathbb{N}}\mathbf{H}_n^f\right|^2, \tag{4}$$

which will be used later on in the analysis.

B. Translation of Lit Structures

We want to manipulate Eq. (4) such that it includes (random) height/length translations of structures in all three spatial directions, $\Delta\mathbf{r} = (\Delta x, \Delta y, \Delta z)$ (see Fig. 2 for a planar example). Considering one of these structures, it is lit by a plane wave with propagation direction $\hat{\mathbf{k}}$ as also indicated in the figure. This means, first of all, that by translating it, we introduce a phase lag, Δp, on the phase that the incoming wave meets the structure with, which can be described as

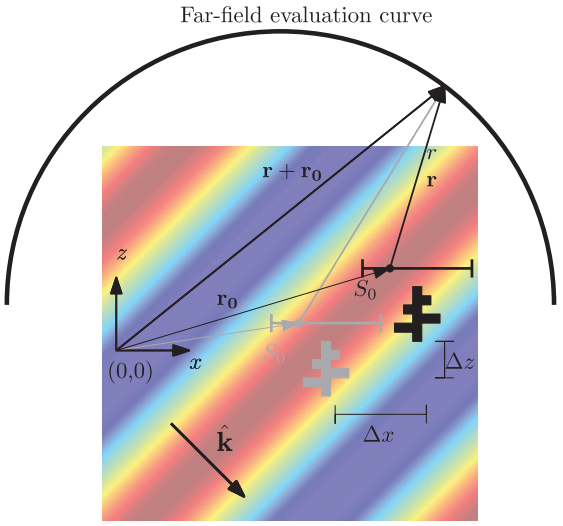


Fig. 2. Geometry of the translation of a structure lit by a plane electromagnetic wave.

$$\Delta p = -\hat{\mathbf{k}}\cdot\Delta\mathbf{r}/\lambda. \tag{5}$$

In relation to the evaluation of the far field, there is also a difference, since the vector $\mathbf{r}_0$ has changed so that it now is represented by

$$\mathbf{r}_0^{\text{new}} = \mathbf{r}_0 + \Delta\mathbf{r}, \tag{6}$$

and this influences the far-field transformation in Eq. (2). We do not need to update the distance, r, since the change in contribution to the distance is negligible, which is in agreement with the assumptions made in the derivation of far-field radiation in the first place [17].

Modifying Eq. (2) then gives

$$\begin{aligned}\mathbf{H}_n^{f,t}(\theta,\phi) &= -e^{-j2\pi\hat{\mathbf{k}}\cdot\Delta\mathbf{r}/\lambda}jk\frac{e^{-jkr}}{4\pi r}\hat{\mathbf{r}} \\ &\quad\times\int_{S_n}\mathbf{J}_e e^{jk\hat{\mathbf{r}}(\theta,\phi)\cdot(\mathbf{r}_o+\Delta\mathbf{r})}\mathrm{d}S_n \\ &= -e^{-jk(\hat{\mathbf{k}}\cdot\Delta\mathbf{r}-\hat{\mathbf{r}}(\theta,\phi)\cdot\Delta\mathbf{r})}jk\frac{e^{-jkr}}{4\pi r}\hat{\mathbf{r}} \\ &\quad\times\int_{S_n}\mathbf{J}_e e^{jk\hat{\mathbf{r}}(\theta,\phi)\cdot\mathbf{r}_o}\mathrm{d}S_n \\ &= e^{-jk(\hat{\mathbf{k}}-\hat{\mathbf{r}}(\theta,\phi))\cdot\Delta\mathbf{r}}\mathbf{H}_n^f(\theta,\phi),\end{aligned} \tag{7}$$

where the superscript t indicates that it is the translated response, and $\mathbf{J}_e$ is still the equivalent surface current of the untranslated structure. This result simply describes the angular-dependent change in phase to the contribution that is seen in an observation point when moving the structure around in a lit domain.

C. Irradiance of Translated Structures

In this paper, we will focus on one basic structure that will exist in several translated instances in the domain (that is, the structure itself will not be pertubed). This structure will be referred to as the unit structure. Due to superposition, there can in principle be several unit structures as long as they do not overlap.

If we assume that we can find the far field radiance of this structure—or at least an adequate approximation of that—we can then sum over this response with the correct translations to find how a system of these structures will reflect light. The validity of the assumption will be discussed in Section 3.G. The far-field radiance has for similar type of problems been calculated by transforming the equivalent surface current simulated over a unit structure using either periodic boundary conditions [18] or radiation boundary conditions [4] depending on the setup.

By having the far-field response for one unit structure, $\mathbf{H}_0^f$, we can then find the total far-field irradiance as

$$E = \frac{1}{2}\eta \left| \sum_n e^{-jk(\hat{\mathbf{k}}-\hat{\mathbf{r}})\cdot\Delta\mathbf{r}_n}\mathbf{H}_0^f \right|^2 = \frac{1}{2}\eta |\mathbf{H}_0^f|^2 \underbrace{\left| \sum_n e^{-jk(\hat{\mathbf{k}}-\hat{\mathbf{r}})\cdot\Delta\mathbf{r}_n} \right|^2}_{=\mathrm{AF}}. \tag{8}$$

In this formulation, the response of the unit structure is isolated such that the total irradiance is just the response of the unit structure multiplied with the magnitude of some function squared. This function we denote the array factor (AF), since it plays the same role as an AF does in antenna theory [17]. Note how the exponent inside the AF has an angular-dependent term, $\hat{\mathbf{r}}(\phi,\theta)$, and a term dependent on the direction of the incoming wave, $\hat{\mathbf{k}}$.

A similar expression appears in [19], but is presented on the basis of scalar diffraction theory.

D. Radiance of Translated Structures

The response of the eye is proportional to radiance and not irradiance [20], meaning that for appearance and color purposes, we need to convert irradiance to radiance.

From [21] it is shown that the irradiance only having a component normal to the observation surface (which is true for a detector in the far field) is related to the radiant intensity, I, by

$$I(\theta,\phi) = r^2 E(\theta,\phi), \tag{9}$$

and also from [21] the relation between radiant intensity and radiance, L, is given as

$$L(\theta,\phi) = \frac{I(\theta,\phi)}{A_s \cos\theta} = \frac{r^2}{A_s \cos\theta} E(\theta,\phi), \tag{10}$$

where A_s is the area of the surface of the lit structure. This means we can write the total radiance for a lit surface with repeated structures as

$$L(\theta,\phi) = \frac{r^2\eta}{2A_s\cos\theta}|\mathbf{H}_0^f(\theta,\phi)|^2 \left| \sum_n e^{-jk(\hat{\mathbf{k}}-\hat{\mathbf{r}}(\theta,\phi))\cdot\Delta\mathbf{r}_n} \right|^2 = \underbrace{\frac{r^2\eta}{2A_0\cos\theta}|\mathbf{H}_0^f(\theta,\phi)|^2}_{=\text{unit response}} N \underbrace{\left| \frac{1}{N}\sum_n e^{-jk(\hat{\mathbf{k}}-\hat{\mathbf{r}}(\theta,\phi))\cdot\Delta\mathbf{r}_n} \right|^2}_{=\mathrm{SAF}}, \tag{11}$$

where A_0 is the surface of $\mathbf{H}_0^f$ and N is the number of summations in the sum, which means that $A_s = N \cdot A_0$. The peculiarity of having N present twice is to scale the last product such that it peaks at unity. Since the AF is now scaled by the number of units, we will refer to it as the scaled array factor (SAF).

The expression in Eq. (11) will be the foundation for all following analyses, as it shows how a prediction of the reflection from the unit structure (this we denote unit response) and a knowledge of the position of its instances (the $\Delta\mathbf{r}_n$'s) can give a complete description of the reflected radiance. Furthermore it decouples the positioning of the structures from the response of the structure, making it possible to analyze each part individually.

E. 1D Version of Formulas

In the rest of this paper, we will consider structures only with variation in the x and z directions and traveling in the x,z plane to keep the examples simple. This means that the results will be invariant in the y plane, and we therefore put $\phi = 90°$ [22]. To indicate this, the following notation will be applied in the rest of this paper:

$$L(\theta) = \frac{r\eta}{2d_0\cos\theta}|\mathbf{H}_0^f(\theta,\phi=90°)|^2 N \times \left| \frac{1}{N}\sum_n e^{-jk(\hat{\mathbf{k}}-\hat{\mathbf{r}}(\theta,\phi=90°))\cdot\Delta\mathbf{r}_n} \right|^2, \tag{12}$$

where d_0 is the length of the unit structure instead of the area. The validity of the SAF concept has been tested using this version of the formula in Appendix A.

F. Interpretation of Array Factor

From Eqs. (11) and (12) it is seen how the total radiance is the product of a unit radiance and the absolute squared SAF. This means that they can be treated independently of each other, and in this paper we will focus only on the SAF and not consider a specific unit response, but instead pose the problem: for a certain unit structure, what role does different randomizations play? The effect of different choices

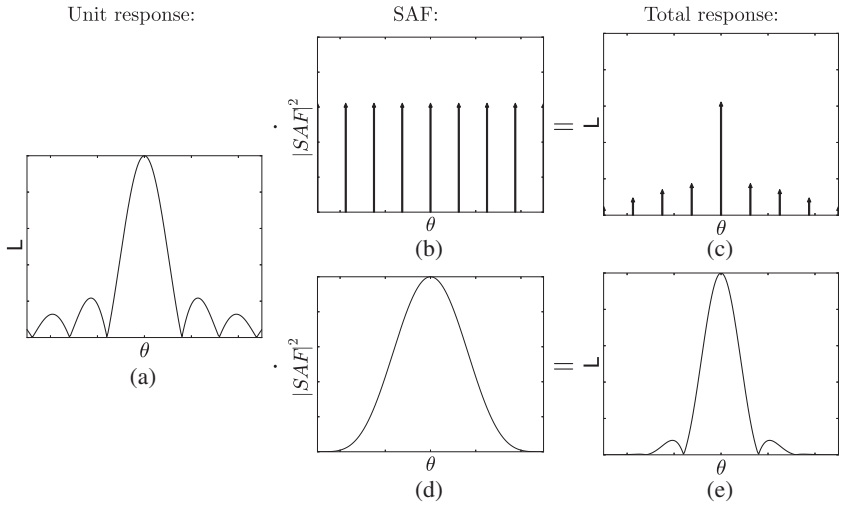


Fig. 3. Example of how a structure with some unit response (a) is influenced by different SAFs. In (b) the SAF for a strictly periodic structure is shown (the comb function), and the effect of combining the structure giving the unit response in (a) with the pattern giving rise to the SAF in (b) is shown in (c). The arrows indicate that all energy is emitted at discrete points. If the SAF is as it is in (d), then the response will end up as seen in (e).

of translations (that is, different SAFs) of the same structure and its influence is exemplified in Fig. 3.

G. Limitations Due to Assumptions

In Sections 3.C and 3.D, we have assumed that all structures on a surface have the same response (except for the translation part), and it is important to clarify when and to what extent these assumptions hold. This is discussed in the following and in Appendix A where a numerical example is presented.

1. *Dependence of Surrounding Geometry*

If a unit structure is simulated sitting in a periodic structure as in Fig. 4(a), and then in reality is sitting in some random structure as in Fig. 4(b), it is clear that the actual far-field response is somewhat different, since the electromagnetic coupling to the neighboring elements has changed. It is therefore necessary that the resulting change in far-field response is small or averages out over many elements, and it is also necessary that the displaced unit structures do not shadow each other. This will in general become less important for large unit structures, since the interelemental coupling in most cases then will be negligible.

It should also be noted that all practical structures have finite sizes, and the unit structures at the edges probably will have another far-field response due to difference in the surrounding geometry. If the surface is large compared to the area occupied by the outer unit structures, this effect should be negligible.

2. *Equivalent Surface Assumption*

We have assumed that there exists a plane surface, S_0, above the unit structure on which we can calculate the equivalent surface current, $\mathbf{J}_e$, and then find the needed unit response from here. Furthermore, we assume that we can calculate the total response from some configuration of unit structures by stitching a plane surface, S, together by these surface currents with a first-order phase-correction term taking their translation into account.

For the above assumptions to work well, we recommend putting the far-field transformation surface as close to the electromagnetic structure as possible for two reasons: the first is to take as much energy as possible into account and thereby catching the behavior in the best way possible; the second is that it will minimize unwanted contributions from the surroundings through the surface (e.g., if there are periodic boundary conditions).

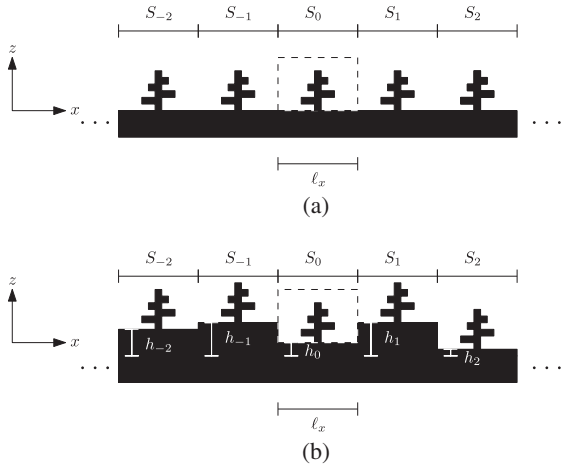


Fig. 4. (a) Strictly periodic structure. (b) The same structure but with per period height translations.

4. Analysis of Random Structures

This section will apply Eq. (12) to different cases—all focusing on explaining the influence of (random) translations of well-defined structures as observed by experiments and simulations in the literature. These examples are chosen to show how randomness can be analyzed and to provide understanding of the effects due to translational randomization.

All results will be in 2D and furthermore with an incoming wave perpendicular to the surface. This means that some important features such as the dependence on the direction of the incoming wave, $\hat{\mathbf{k}}$, will not be discussed in this paper.

A. No Randomness: A Special Case Called Diffraction

To introduce the concept and confirm its validity for a known (but trivial) case, the first example will be the rather simple case of a repeated structure with no randomization.

Consider an infinitely periodically repeated unit structure within a domain [see Fig. 4(a)]. If this structure is lit with a normal incident plane wave, then we have that

$$\hat{\mathbf{k}} = (0, 0, -1), \qquad \Delta\mathbf{r}_n = (n\ell_x, 0, 0), \qquad n \in \mathbb{Z}, \quad (13)$$

where ℓ_x is the period with which the structure is translated. The SAF for this structure then becomes

$$\mathrm{SAF}(\theta) = \frac{1}{N}\sum_{n\in\mathbb{Z}} e^{-jk(\hat{\mathbf{k}}-\hat{\mathbf{r}}(\theta))\cdot\Delta\mathbf{r}_n} = \frac{1}{N}\sum_{n\in\mathbb{Z}} e^{jkn\ell_x \sin\theta}. \quad (14)$$

For infinite repetition, the limit of the summation inside the norm can be expressed as [23]

$$\sum_{n\in\mathbb{Z}} e^{j2\pi n\frac{\ell_x}{\lambda}\sin\theta} = \sum_{n\in Z}\delta\left(\frac{\ell_x}{\lambda}\sin\theta - n\right) =: \mathrm{comb}\left(\frac{\ell_x}{\lambda}\sin\theta\right), \quad (15)$$

where δ is the Dirac delta function, and the name *comb* is given due to the function's resemblance with a comb [see Fig. 3(b)]. It is also referred to as the *shah* function. Since $\delta(x) = 0, \quad \forall\, x \in \mathbb{R}\backslash\{0\}$, reflection from this structure will only appear when

$$\frac{\ell_x}{\lambda}\sin\theta = m \in \mathbb{Z} \quad (16)$$

and have strong intensity. The relation in Eq. (16) is called the grating equation and m is normally referred to as the mode number. It is a well-known result, explaining, e.g., the rainbow effect seen when observing a compact disc (CD), which consists of equally spaced grooves used to store the data. This behavior is referred to as diffraction and is seen to be explained by Eq. (12) as expected.

B. Random Height Variation of Repeated Structure

Consider the same setup as before with an infinitely repeated structure [see Fig. 4(a)]. In the analysis of the coloration of the Morpho butterfly, numerical simulations in [13] shows that by adding a random height variation to each unit, the total response will start to resemble the unit response with overlayed high-frequency ripples. To analyze this, we apply a random height translation to each unit drawn from a uniform distribution with values between 0 and $\lambda_{\max}$, where $\lambda_{\max}$ is the longest wavelength in the analysis. The final structure will then look like the one seen in Fig. 4(b). The incoming wave is still normal to the surface, so the components for the SAF of this system are

$$\hat{\mathbf{k}} = (0, 0, -1), \qquad \Delta\mathbf{r}_n = (n\ell_x, 0, \ell_{z,n}), \qquad n \in \mathbb{Z}, \quad (17)$$

where ℓ_x is the period and $\ell_{z,0}, \ell_{z,1}, \ell_{z,-1}, \ell_{z,2}\ldots$ is a sequence of numbers drawn from a uniform distribution with values between 0 and $\lambda_{\max}$. This means that the SAF now takes the form

$$\begin{aligned}\mathrm{SAF}(\theta) &= \frac{1}{N}\sum_{n\in\mathbb{Z}} e^{-jk(n\ell_x \sin\theta - (\cos\theta+1)\ell_{z,n})} \\ &= \frac{1}{N}\sum_{n\in\mathbb{Z}} e^{-jkn\ell_x \sin\theta} e^{jk(\cos\theta+1)\ell_{z,n}}, \end{aligned} \quad (18)$$

where we see that the first product in the summation comes from the periodic translation, and the second from the random height translation. The minus sign in the second product indicates that positive height displacements reflect the phase earlier, and the angle dependent $1 + \cos\theta$ can be interpreted geometrically as the extra added distance the wave has to travel. That is, for specular reflection where $\theta = 0$, the wave will also travel the same phase less than it did when reaching the structure, but for other values of θ, it will travel a bit longer, and therefore not as much negative phase lag will be removed.

A typical response of Eq. (18) is seen in Fig. 5(a) for a finite number of structures ($N = 100$). It is seen how $\ell_{z,n}$ should vary between zero and half a wavelength to get a SAF where no diffraction pattern is present and also that if there is no randomization, then the SAF has sharp intensities peaking in the grating modes that can be calculated from Eq. (16). The finite number of repetitions spreads out the intensities from just a single angle to a small angular area. It is also seen that there is a gradual change from pure diffraction to no diffraction with the change of wavelength. For larger maximum values of $\ell_{z,n}$, the diffraction is in general weakly present if present at all.

A good way of describing the above observations intuitively is by considering what phase is most probable to be observed in a certain far-field direction, and if there is no preferred phase then there is no possibility of interference, whereas if there is a preference of a phase, then that will give rise to interference effects (e.g., if the phase is varying in a smaller interval than $\theta \in [0°, 360°]$).

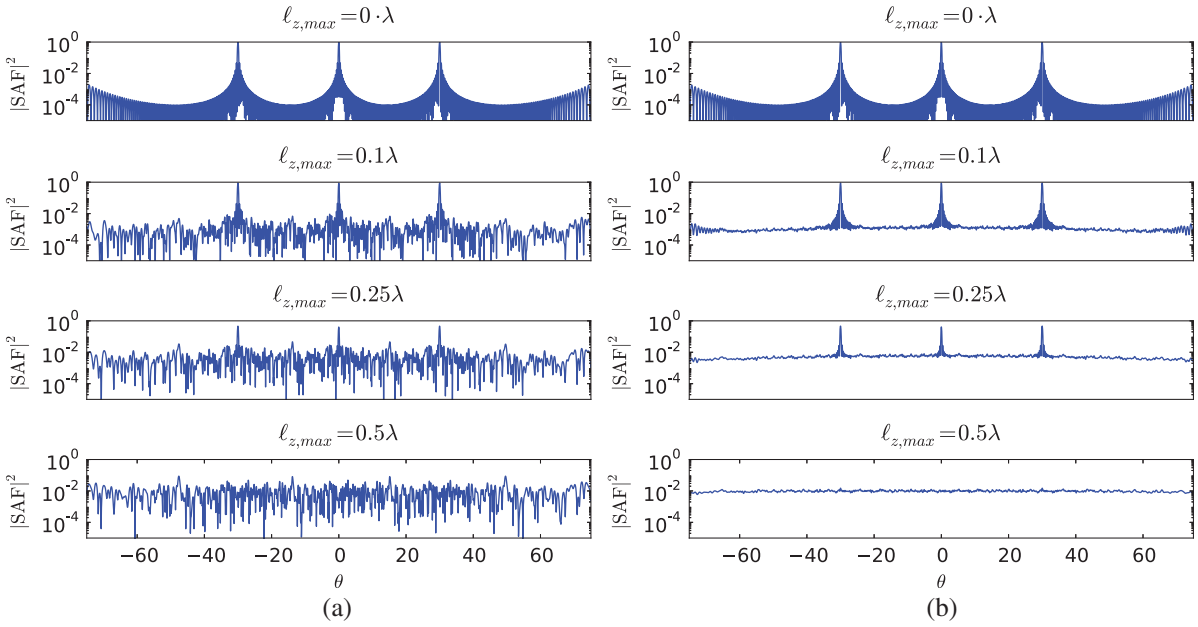


Fig. 5. (a) Numerical calculation of Eq. (18) with $\ell_{z,n}$ picked uniformly from $[0, \ell_{z,\max}]$ and $N = 100$ repetitions and a periodicity in x of $\ell_x = 2\lambda$. By changing ℓ_x the SAF would either be dilated or constricted such that the grating modes still match Eq. (16). (b) 100 averages of the setup in (a) using Eq. (19).

These results can be confirmed by [15] where they report a clearly visible diffraction pattern for small variation and show that larger variations are needed to minimize the effect. Furthermore it can be seen in Fig. 4(c) in [15] how the 380 nm still shows the first-order diffraction at $\theta = \arcsin(380/400) \approx 72°$, and otherwise a random pattern with strong "noise" and a zero-order mode is still present, since the variations are smaller than a wavelength, and furthermore how longer wavelengths (as just shown) gives a larger contribution to mode 0 (when the relative reflection from Fig. 3 in [15] is taken into account). Furthermore, the same analysis reveals the governing effect of Fig. 6 in [16] in the explanation of the results in [16].

1. *Added Incoherence*

In reality, the reflected spectrum has less ripples than in Fig. 5(a). In [15,24], this has been taken into account by averaging the irradiance of a large ensemble to add the effect of incoherence. The explanation for doing so is that the phases of two uncorrelated waves on average will neither add destructively nor constructively, and it is therefore possible just to sum their powers. That is,

$$|\mathrm{SAF}(\theta)|^2 = \left\langle \left| \frac{1}{N} \sum_{n \in \mathbb{Z}} e^{-jkn\ell_x \sin\theta} e^{jk(\cos\theta + 1)\ell_{z,n}} \right|^2 \right\rangle, \quad (19)$$

where $\langle \cdot \rangle$ for this equation indicates that the average will be taken of the SAF for many different seeds of $\ell_{z,n}$. By doing this averaging, we end up with the result in Fig. 5(b), where the "noise ripples" are now much smaller. This is in good agreement with the results from [15], and, for an even higher number of averages, the ripples become even smaller.

C. Triangular Height Distribution

In [13], the height displacements are chosen from a triangular distribution. Using Eq. (18), where $\ell_{z,n}$ is now drawn from a triangular distribution ranging from 0 to $\ell_{z,\max}$ with the triangular peak placed in $\ell_{z,\max}/2$, gives the results presented in Fig. 6. This result has more visible diffraction than for the uniform samples. In particular it is seen how they are still present for $\ell_{z,\max} = 1/2\lambda$. Investigating the Ph.D. thesis on which the article is based [25] reveals that the interval for the triangular distribution is also double the size of the uniform distribution it is compared with. Doing the analysis with a broader distribution gives the results in Fig. 7, and from here it is seen that the modes actually are better suppressed, which explains the choice of distribution in [13]. It was chosen to plot the incoherent/averaged version here, since it is then easier to compare with Fig. 5(b).

D. Binary Randomized Heights

In [14], a blue surface without specular reflections is produced to mimic the behavior of the Morpho butterfly wing. For easy fabrication purposes, the randomization of the unit structure is made up by only two heights; that is, $\ell_{z,n}$ in Eq. (18) can take on two values: 0 and $\ell_{z,\max}$. Repeating the setup from before, but now with this binary randomness, we obtain the results presented in Fig. 8. Here it is seen how a height difference of a quarter of a wavelength

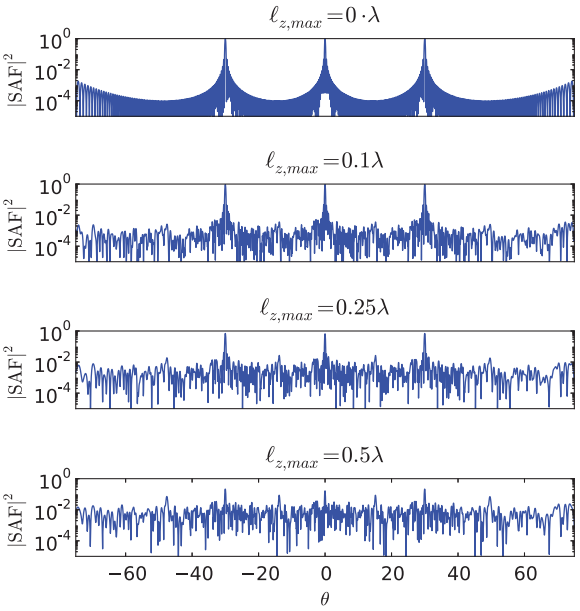


Fig. 6. Results for the same setup as in Fig. 5(a), but now with $\ell_{z,n}$ taken from a *triangular* distribution.

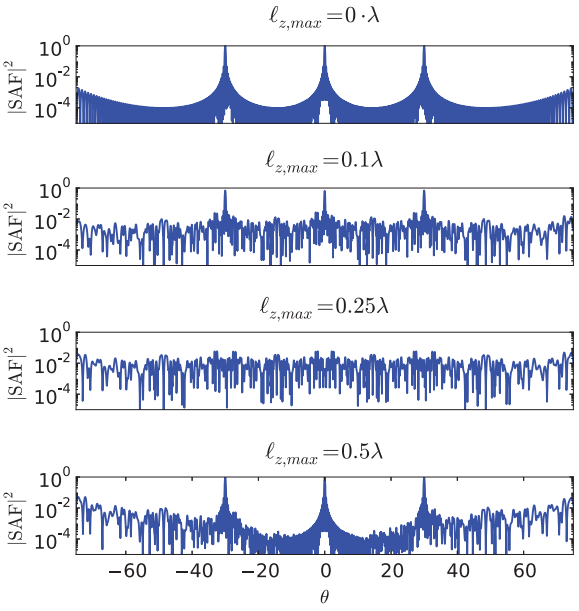


Fig. 8. Results for the same setup as in Fig. 5(a), but now with $\ell_{z,n}$ only taking the values 0 and $\ell_{z,\max}$.

at the target frequency is needed to get a flat $|\mathrm{SAF}|^2$. This is in perfect agreement with the height chosen in [14]. Choosing a height difference on half a wavelength gives strong grating modes. The reason is that for a height difference of $1/4\lambda$, the reflected light is sent back 180° out of phase and therefore interferes destructively, leaving no preferred directions for the reflected light, whereas if the light travels $1/2\lambda$ extra back and forth it corresponds to a 360°

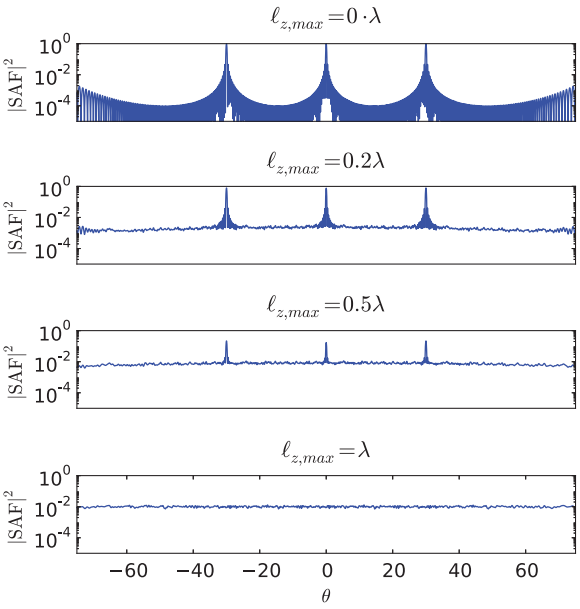


Fig. 7. 100 averages of the same setup as Fig. 6(a), but for double the interval width of the distribution used in Fig. 6(a).

phase shift, which gives constructive interference and modes as seen in Fig. 5(a). Since the reflected light travels a bit longer than $1/2\lambda$ when reflected back in off-specular directions (taken care of in the equation by the cos θ term), the effect of constructive interference wears off more and more for increasing angles, which is also seen on the plot when comparing $\ell_{z,\max} = 0$ with $\ell_{z,\max} = 1/2\lambda$.

This example shows the need for analyzing specific randomizations, as here $\ell_{z,\max}$ has to be chosen differently, and since the effect only can be expected to work well for a smaller range of wavelengths than in the uniform randomness case. In the design in [14], the wavelengths far from the blue region around 440 nm are suppressed by the unit structure, thus hiding the diffraction effects that would otherwise have been present.

E. In-Plane Translated Elements

In [15], in-plane movement is considered; that is, movement in the x direction of a repeated structure. For this setup, the following definitions for the SAF are needed:

$$\hat{\mathbf{k}} = (0, 0, -1), \qquad \Delta\mathbf{r}_n = (n\ell_x + \Delta x_n, 0, 0), \qquad n \in \mathbb{Z}, \tag{20}$$

where the Δx_n's is a sequence of random variables drawn uniformly from the range $[-\Delta x_{\max}/2, \Delta x_{\max}/2]$. This gives rise to the following SAF:

$$\begin{aligned} \mathrm{SAF}(\theta) &= \frac{1}{N}\sum_{n\in\mathbb{Z}} e^{-jk(n\ell_x + \Delta x_n)\sin\theta} \\ &= \frac{1}{N}\sum_{n\in\mathbb{Z}} e^{-jkn\ell_x \sin\theta} e^{-jk\Delta x_n \sin\theta}. \end{aligned} \tag{21}$$

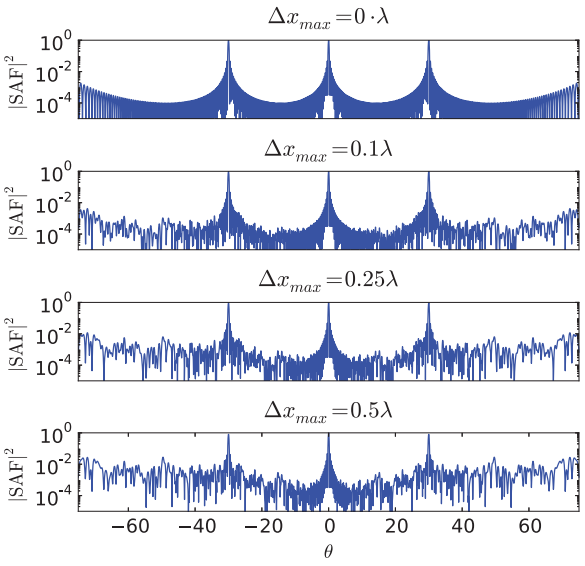


Fig. 9. Results for the same setup as in Fig. 5(a), but with in-plane movement as specified in Eq. (20).

By simulating Eq. (21), we get the results shown in Fig. 9. Comparing this with Eq. (21), it can be seen that in-plane translation will never affect the specular mode response (mode 0) since $\sin\theta = 0$ in that direction, which means that $\mathrm{SAF}(0) = 1$ no matter the randomization. For larger angles, the effect will be more and more prominent, though, since $\sin\theta$ is larger and the preferred phase will be less prominent in the phase distribution. This is a huge limitation with respect to creating an effect with no visible diffraction, but could add some randomness to large angles. This is also what is observed in [15].

It should be noted that the displacements in Fig. 9 are quite large, and it would require a structure with lots of air in between as for the structure in this example not to overlap or couple significantly.

5. Color Representation of SAFs

To give an idea of the interpretation of the obtained results in terms of color effects for a surface, this section will present some of the SAFs converted to RGB values [26] for given sizes. The conversion used can be found in [27]. The examples used all have a period of $\ell_x = 2$ μm, and differ by having (1) a random height variation drawn from a uniform distribution between 0 and 110 nm; (2) a random height variation drawn from a uniform distribution between 0 and 220 nm; (3) a random height variation drawn from a uniform distribution between 0 and 1500 nm; (4) and a uniform binary height displacement with the values 0 and 110 nm, which corresponds to the parameters chosen in [14]. The results are presented in Fig. 10.

What can be seen from these plots is that if a uniform random distribution is chosen, and we are designing for a color around a certain wavelength, λ, then $1/2\lambda$ should be chosen as the upper limit for the

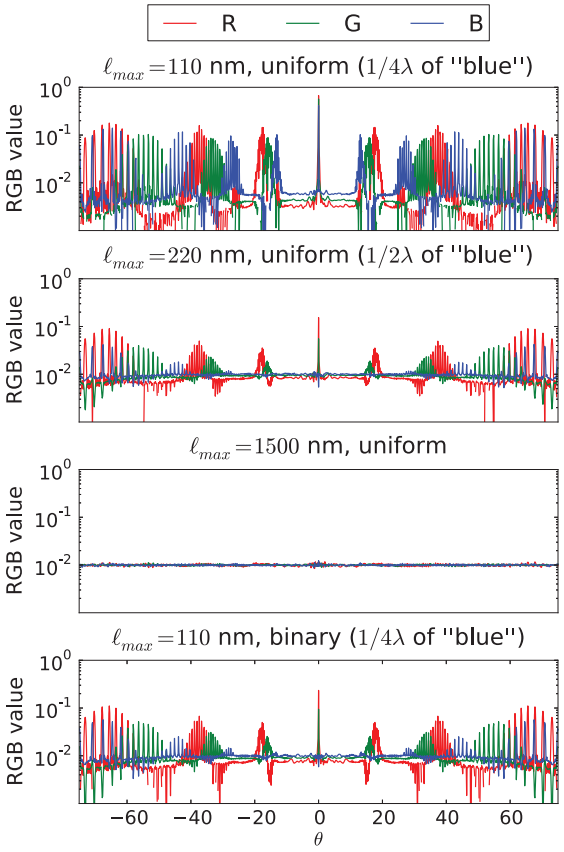


Fig. 10. Color representation of different SAFs. All unit structures are repeated with a period of 2 μm, and incoherence has been taken into account by averaging over 200 samples. The oscillatory effect seen at large angles is due to the fact that the angular spacing between wavelengths gets larger. Notice how the response for blue is flat for all plots except the uniform random distribution only going to 110 nm.

uniform distribution (blue has a wavelength around 440 nm), whereas a binary random distribution will require only $1/4\lambda$, and, as shown earlier, choosing $1/2\lambda$ will give a strong diffraction pattern. Furthermore it is seen that choosing a large upper limit of the uniform distribution makes a good choice for giving a flat distribution of all colors. This could have some practical issues, though, since the translation may be large compared to the unit structure and possibly violate the assumptions stated earlier. All in all, this suggests that when designing random patterns for a given color, it may be beneficial to use binary randomness if possible, since this gives the smallest translation of the structures and in many cases will be easier to produce by, e.g., an etching process as in [14]. Not all colors can be represented by just one relatively small band in the visible range such as, e.g., magenta or white, and if that is not possible other means are needed. This could, for example, be large uniform randomness or possibly the concept of binary randomness but expanded to more levels,

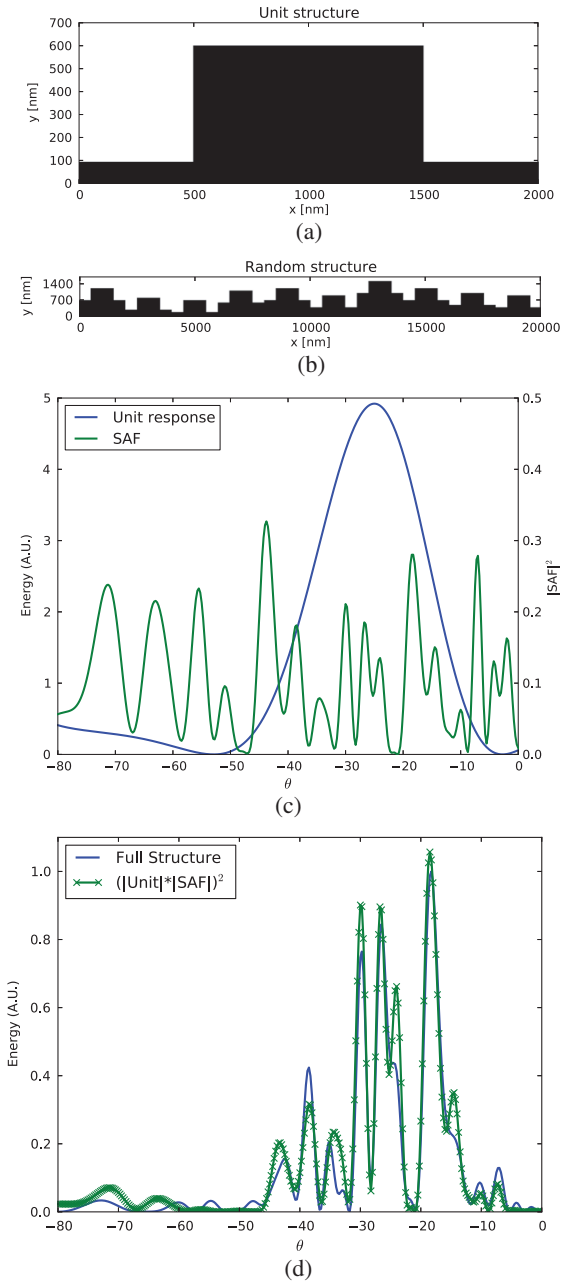


Fig. 11. Verification of the proposed method by comparison to a full-wave simulation for the E_z polarized case (c.f. [26]). (a) Unit structure with black representing material having $\varepsilon_r = 3.5$ and white being air. (b) Ten random height translations of this structure with the heights chosen as described in the text. Both structures have periodic boundary conditions. The far-field transformation for (a) and (b) are calculated on the lines $y = 700$ nm, for $x \in [0, 2000]$ nm and $y = 1500$ nm, for $x \in [0, 20000]$ nm, respectively. (c) Unit response for (a) simulated at $\lambda = 750$ nm as well as the SAF calculated from Eq. (22). (d) Comparison of the approximation and full simulation reference.

which are chosen such that they provide destructive interference at several wavelengths. The method will be dependent on the target color, and further investigations are therefore left for the concrete cases that a scientist or designer may have.

6. Conclusion

A method for analyzing translatorial randomness of repeated structures has been presented, and it has been shown how existing observations in the literature can be explained from this. The results presented are intented to push forward the understanding of these randomization phenomena; for example, it explains why height randomization removes diffraction patterns much better than in-plane randomization and how to test the effect of different height distributions. This saves time and gives more insight in the analysis of random phenomena compared to the more expensive full-wave simulation of repeated structures as has been seen earlier in the investigation of random effects. The method also makes it possible to apply a systematic approach for choosing randomization characteristics when designing surfaces with new color effects based on a unit structure.

Appendix A: Verification Example

To verify the presented method, a unit structure as seen in Fig. 11(a) has been repeated 10 times with a random height translation and lit by a wave having an incident angle of 25° to the normal of the surface. The response has then been calculated using Eq. (12) and compared to a full-wave simulation of the whole structure [see Fig. 11(b)]. This means that

$$\hat{\mathbf{k}} = (\sin\theta_i, 0, \cos\theta_i), \qquad \Delta\mathbf{r}_n = (n\ell_x, 0, \ell_{z,n}), \quad \text{(A1)}$$

where $\theta_i = 180° + 25°$, $\ell_x = 2$ μm, and $\{\ell_{z,n}\} = \{0.6, 0.2, 0.1, 0.5, 0.6, 0.3, 0.9, 0.6, 0.4, 0.3\}$ μm. The results of the simulation are presented in Fig. 11. By visual inspection, a fairly good match can be found; note in particular how most peak positions are caused purely by the translatorial properties and how the intensity distribution is dominated by the unit structure response.

The author would like to thank Professors Olav Breinbjerg and Ole Sigmund for discussions and excellent feedback regarding this paper. The author is grateful for the support from the Danish National Technology Foundation through the ODAAS project.

References

1. R. Hooke, Micrographia, http://www.gutenberg.org (1665).
2. A. R. Parker, "515 million years of structural colour," J. Opt. A **2**, R15–R28 (2000).
3. P. Vukusic, J. R. Sambles, C. R. Lawrence, and R. J. Wootton, "Quantified interference and diffraction in single Morpho butterfly scales," Proc. R. Soc. B **266**, 1403–1411 (1999).
4. N. Okada, D. Zhu, D. Cai, J. B. Cole, M. Kambe, and S. Kinoshita, "Rendering Morpho butterflies based on high accuracy nano-optical simulation," J. Opt. **42**, 25–36 (2013).
5. C. W. Mason, "Structural colors in insects. II," J. Phys. Chem., **31**, 321–354 (1927).

6. W. Lippert and K. Gentil, "Über Lamellare Feinstrukturen bei den Schillerschuppen der Schmetterlinge vom Urania- und Morpho-typ Z," Morph. Ökol. Tiere **48**, 115–122 (1959).
7. A. R. Parker, T. Lenau, and A. Saito, "Biomimetics of optical nanostructures," in *Biomimetics in Photonics* (CRC Press, 2012), pp. 55–115.
8. S. Kinoshita, D. Zhu, and A. Saito, "Modeling and simulation of structural colors," in *Biomimetics in Photonics* (CRC Press, 2012), pp. 191–242.
9. P. Licinio, "Diffraction by disordered gratings and the DebyeWaller effect," Am. J. Phys. **67**, 1013–1016 (1999).
10. J. M. Rico-García and L. M. Sanchez-Brea, "Binary gratings with random heights," Appl. Opt. **48**, 3062–3069 (2009).
11. T. Buß, J. Teisseire, and S. Mazoyer, "Controlled angular redirection of light via nanoimprinted disordered gratings," Appl. Opt. **52**, 709–716 (2013).
12. F. Pratesi, M. Burresi, F. Riboli, K. Vynck, and D. S. Wiersma, "Disordered photonic structures for light harvesting in solar cells," Opt. Express **21**, A460–A468 (2013).
13. R. T. Lee and G. S. Smith, "Detailed electromagnetic simulation for the structural color of butterfly wings," Appl. Opt. **48**, 4177–4190 (2009).
14. A. Saito, Y. Miyamura, Y. Ishikawa, J. Murase, M. Akai-Kasaya, and Y. Kuwahara, "Reproduction, mass-production and control of the Morpho-butterfly's blue," Ad. Fabric. Technol. Micro/Nano Optics and Photonics II **7205**, 720506 (2009).
15. A. Saito, M. Yonezawa, J. Murase, S. Juodkazis, V. Mizeikis, M. Akai-Kasaya, and Y. Kuwahara, "Numerical analysis on the optical role of nano-randomness on the Morpho butterfly's scale," J. Nanosci. Nanotech. **11**, 2785–2792 (2011).
16. M. A. Steindorfer and V. Schmidt, "Detailed simulation of structural color generation inspired by the Morpho butterfly," Opt. Express **20**, 21485–21494 (2012).
17. C. A. Balanis, *Advanced Engineering Electromagnetics*, 2nd ed. (Wiley, 2012).
18. M. Zhou, S. B. Sørensen, E. Jørgensen, P. Meincke, O. S. Kim, and O. Breinbjerg, "An accurate technique for calculation of radiation from printed reflect arrays," IEEE Antennas and Wireless Propagation Lett. **10**, 1081–1084 (2011).
19. D. Zhu, S. Kinoshita, D. Cai, and J. Cole, "Investigation of structural colors in Morpho butterflies using the nonstandard-finite-difference time-domain method: effects of alternately stacked shelves and ridge density," Phys. Rev. E **80**, 051924 (2009).
20. P. Dutré, K. Bala, and P. Bekaert, *Advanced Global Illumination* (A K Peters, 2006).
21. J. E. Harvey, C. L. Vernold, A. Krywonos, and P. L. Thompson, "Diffracted radiance: a fundamental quantity in nonparaxial scalar diffraction theory," Appl. Opt. **38**, 6469–6481 (1999).
22. T. Antonakakis, F. Bada, A. Belkhir, K. Cherednichenko, S. Cooper, R. Craster, G. Demesy, J. DeSanto, G. Granet, B. Gralak, S. Guenneau, D. Maystre, A. Nicolet, B. Stout, F. Zolla, and B. Vial, *Gratings: Theory and Numeric Applications*, 1st ed. (Presses Eniversitaires de Provence, 2012).
23. R. N. Bracewell, *The Fourier Transform and its Applications*, 3rd ed. (McGraw Hill, 2000).
24. S. Kinoshita, S. Yoshioka, Y. Fujii, and N. Okamoto, "Photophysics of structural color in the Morpho butterflies," Forma, **17**, 103–121 (2002).
25. R. T. Lee, "A novel method for incorporating periodic boundaries into the FDTD method and the application to the study of structural color of insects," Ph.D. thesis (Georgia Institute of Technology, 2009).
26. R. S. Berns, F. W. Billmeyer, and M. Saltzman, *Billmeyer and Saltzman's Principles of Color Technology* (Wiley-Interscience Publication, 2000).
27. J. Andkjær, V. E. Johansen, K. S. Friis, and O. Sigmund, "Inverse design of nanostructured surfaces for color effects," J. Opt. Soc. Am. B **31**, 164–174 (2014).

Publication P4

Designing visual appearance using a structured surface

Designing visual appearance using a structured surface

VILLADS EGEDE JOHANSEN[1,*], LASSE HØJLUND THAMDRUP[2], KRISTIAN SMISTRUP[2], THEODOR NIELSEN[2], OLE SIGMUND[1], AND PETER VUKUSIC[3]

[1] Technical University of Denmark, Department of Mechanical Engineering, 2800 Kgs. Lyngby,Denmark
[2] NIL Technology ApS, Diplomvej 381, 2800 Kgs. Lyngby, Denmark
[3] School of Physics, University of Exeter, EX4 4QL, UK
[*] Corresponding author: vejo@mek.dtu.dk

Compiled January 23, 2015

We present an approach for designing nanostructured surfaces with prescribed visual appearances, starting at design analysis and ending with a fabricated sample. The method is applied to a silicon wafer structured using deep ultraviolet lithography and dry etching and includes preliminary design followed by numerical and experimental verification. The approach comprises verifying all design and fabrication steps required to produce a desired appearance. We expect that the procedure in the future will yield structurally colored surfaces with appealing prescribed visual appearances.

OCIS codes: *(330.1690) Color; (050.1960) Diffraction theory; (240.6700) Surfaces; (290.0290) Scattering theory; (330.7326) Visual optics, modeling*

http://dx.doi.org/10.1364/optica.XX.XXXXXX

1. INTRODUCTION

Advances in nanotechnology are improving possibilities for designing structural colors. Recently a range of different techniques for creating synthetic structurally colored systems have been reported in literature. The methods each utilize very different production techniques ranging from self-assembly [1–4] to deposition, growth, embossing and etching techniques [5–14]. The nanofabrication methods are often highly inventive with different limitations on geometries and materials thus also on obtainable colors and angle dependence. The governing effects for color control are (layered) interference effects possibly combined with randomization [1, 5, 7–9, 11, 12, 14], structural (particle) scattering [2–4, 6] and surface plasmon effects [10, 13]. The mentioned approaches – which by no means provide an exhaustive list– all have their limitations due to manufacturability. For example, good color selectivity but little control of angular reflection or expensive fabrication and/or design procedures. We therefore find that there still is room to discover methods that are: reproducible on an industrial scale; cheap; applicable to different surface types and curvatures; environmentally friendly; and controlled tunable.

Another important point is that all approaches considering surface coloration [1, 3–6, 8, 10, 13, 14] except two [7, 12] base their color selectivity and/or angular dependence on adding new materials to the surface. The two exceptions both have interference-like reflections, not smooth angular variations in color. Exploring the possibilities of surface manipulation techniques without addition of secondary materials to obtain a smooth, appealing color effect is one method to address many of the issues above.

In this paper we report results of our initial studies towards producing structural colors, that will address all the issues above, except tunability, by considering large batch manufacturable (preferably injection moldable) structured surfaces. The design freedom is currently limited by manufacturing constraints and tolerances for the choice of production processes. Despite this we are able to demonstrate here how to navigate the design and manufacturing process from specifications to a fabricated surface. We see this as a valuable step towards widespread use of structural color design in an industrial setting.

2. DESIGN APPROACH OUTLINE

The goal, when designing a system to exhibit structural colors, is to identify a structure with a desired visual response, cf. Figure 1(a). The design of sub-micrometer features on a surface that extends millimeters or more is an overwhelming task. For this

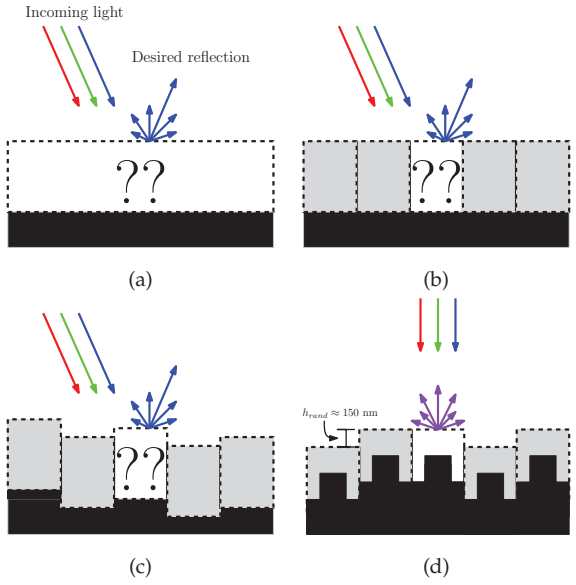


Fig. 1. (color online) (a) A general structural color problem: find a design that gives a prescribed reflection – e.g. a blue distribution – for incoming white light. (b) Due to many orders of magnitude difference in the analysis (ranging from a centimeter sized surface to sub-micrometer electromagnetic waves) the analysis has to be divided into subparts, which is done here by assuming periodic repetition of the same structure. (c) Since a pure periodic structure gives rise to unwanted diffraction effects, this has to be avoided by also designing a way of "randomizing" the unit structure – here it is done by height translations. (d) In order to verify the approach in (a-c) a test structure is produced to see how measurements fit analysis.

reason, if no systematic approach is used, the problem has to be divided into manageable components. An obvious choice that makes the computational task of predicting color reflection more accessible is to assume a form of structure that is repeated across the surface, such as is shown in Figure 1(b). A regular structure with a fixed period is unfit for our purpose due to interference effects [15]. To avoid this we introduce height randomization in these repeated units [15, 16]. The dominant effect of such an approach can be approximated analytically [15] and verified numerically before fabrication. This approach is represented in Figure 1(c).

Using this technique means that the design challenge is broken down into two (decoupled) components: the design of a unit structure, and the design of its randomization. In earlier works[15, 17] we have shown how to design such unit structures for various 2D design goals, and we have also described how to incorporate the effect of randomization [15]. Production of such samples is challenging and to test this approach a sample is fabricated using deep ultraviolet lithography (DUV) and dry etching. This will be described in Section 5. This means that only three levels of height can be used in designing the structure and its randomization, meaning that the final design will appear as shown in Figure 1(d). A flow chart of this general design procedure applied to our study is depicted in Figure 2.

The following four sections describes the individual design steps in Figure 2.

3. STEP 1: STRUCTURE DESIGN

Our previous works [15, 17–19] focused on how to model and design dielectric structures for prescribed color responses in a systematic way. In this paper we focus on the sample production and verification steps and have therefore chosen to use a 50% duty cycle (half-pitch) line grating with a height of 215 nm as the unit structure for our structurally colored surface. This structure can be produced using traditional methods and equipment at our disposal. The choice of height gives several distinct specular reflection peaks and dips in the visible spectrum. In particular, simulations will show that both a strong dip and peak exist below a wavelength of 500 nm, thus making it possible to investigate how well energy either can be kept in – or transferred from – specular reflection due to overlaid randomization. Since line gratings are invariant in depth, it means that if we chose a randomization with the same invariance properties, the reflection of light will be confined to a plane, making experimental measurements more straight-forward. A duty cycle of 50% also gives the largest color contrast with respect to specular reflection [19, Fig. 7(a)]. A binary random pattern with a height of 150 nm is chosen for the randomization, which maximizes scattering around 600 nm [15].

The unit structure period is chosen to be 500 nm so that shorter visible wavelengths will diffract and larger visible wavelengths will not, for normal incidence. This makes it possible to capture different types of behavior in one measurement at normal incidence. It also complies with the choice of line grating height, since light from the strong dip and peak will be diffracted.

To predict the reflection, a one unit structure was simulated for unpolarized light (power average of E_z, H_z) using the finite element method (FEM) with periodic boundary conditions as described in earlier work [17] and afterwards post-processed by incorporating the random effect as described in [19]. The optical properties of silicon were taken from literature [20]. This means that the scattered unit far field radiance L_u was calculated first and was subsequently multiplied by what we called the scaled array factor (SAF)[19]. The total radiance L can be described as

$$L = L_u \underbrace{\frac{1}{M}\sum_{m=1}^{M}\left|\frac{1}{N}\sum_{n=1}^{N} e^{-ik(n\ell_x \sin\theta + \Delta z_n(\cos\theta+1))}\right|^2}_{SAF}, \qquad (1)$$

where θ is the reflection angle, i is the imaginary unit, ℓ_x is the length of one period, Δz_n is a uniform random number taking on either 0 or 150 nm, N and M are loosely connected to the coherence of the incoming light [16, 19]. The resulting reflection of the unit structure is shown in Figure 6(a) and the SAF is shown in Figure 6(b). By multiplying the results they produce the total radiance shown in Figure 6(c).

4. STEP 2: DESIGN VERIFICATION

The approximation described in Step 1 provides an efficient route for analysis which is often needed in the design phase. To verify the approximation of the final design, full FEM simulations were used (with the same number of averages and choices of Δz_n). The result is presented in Figure 6(d) and takes several days to generate on an 8 core machine, since each model contains $5 \cdot 10^5$ elements and an average of approximately $M = 100$

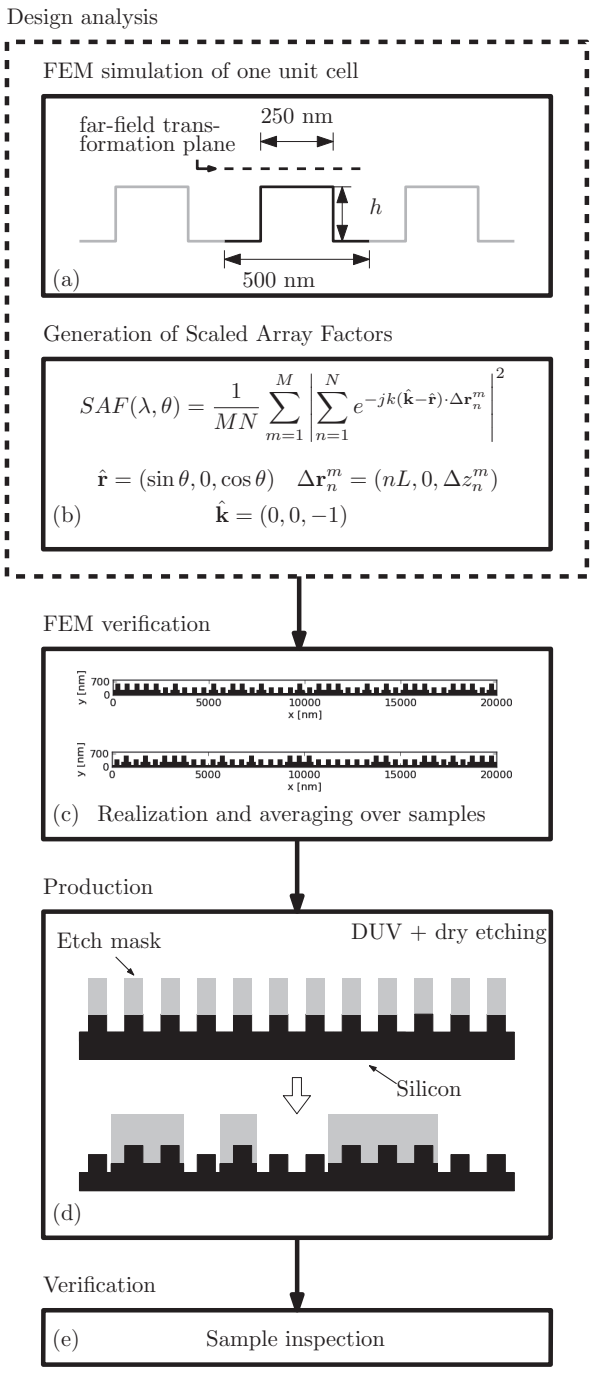


Fig. 2. Design workflow used to create a structurally colored surface – in this case by using DUV and dry etching. (a) First a unit structure is defined and analyzed; (b) at the same time the effect of the scaled array factor (SAF) is calculated and taken into account, (c) this design is then verified using a full FEM model, (d) this then gives the dimensions from which to produce a sample. (e) Lastly the result is verified.

models is needed for a meaningful response. This is why a full simulation is not feasible for the design phase.

The overall behavior is seen to be similar to the behavior of the approximate model. The most notable differences are that: the very low intensity areas and reflection minima are less pronounced; the peak intensities are in general lower; and the first order diffraction modes seems to have shifted in intensity. We expect even better fits for structures with more smooth angular reflection spectra.

5. STEP 3: SAMPLE PRODUCTION

The sample was produced on a single-side polished, 100 mm diameter, silicon wafer with a thickness of 525 μm. A sketch of the process steps is presented in Figure 2(d). The optimal DUV dose for the resist (JSR-M230Y from JSR Micro, 350 nm thickness) exposure of the underlying periodic grating was found to be 220 J/m^2 to ensure a 50% duty cycle after pattern transfer into silicon by advanced silicon etching (STS MESC Multiplex ICP Advanced Silicon Etcher). The second DUV exposure was performed in JSR-M35G resist (from JSR Micro), thickness 750 nm, at a DUV dose of 320 J/m^2 to ensure complete coverage of the periodic grating. The design of the DUV reticle did not incorporate dose compensation, so we aimed for maximum pattern replication fidelity of structures with the smallest line width (i.e. 500 nm). The dry etching of the random grating was performed using an inductive coupled plasma etch (SPTS ICP Metal Etcher) to ensure a high precision in the etch depth. Afterwards the sample wafer was cleaned using oxygen plasma and Piranha solution.

6. STEP 4: SAMPLE VERIFICATION

A. Geometry inspection

The geometry of the produced sample was inspected using scanning electron microscopy (SEM) and atomic force microscopy (AFM). A selection of SEM images are presented in Figure 3: Figure 3(a) shows one corner of the sample and it is possible to see the misalignment between the two reticles that were been used. The misalignment is almost purely translational, with less than 10^{-3} degrees of rotation with respect to each other. This means that the two patterns are experiencing a shift of around 50 nm per centimeter. Since our samples are 1×1 cm^2, no gradual colour change was observed due to this effect. Misalignment is expected to affect reflection properties for this test geometry over larger areas. See Media 1 for a representation of the effect of misalignment. At the same time we do not believe that the effect will be observable under normal lighting conditions – in particular for a geometry optimized for an appealing color effect, where the randomization will be in 2D [5]. Figure 3(b) shows that the period of the grating is approximately 505 nm, and the trenches are measured to be around 243 nm. Figure 3(b) also shows that the difference between the trenches that have been etched further down by the randomization process and those that have not, is small. Figure 3(c) shows a more distinct difference from the predicted design, namely, the width of the trenches of the random overlay have been etched wider than planned ($\sim$ 30 nm on each side in the area without the underlying grating). These trenches can also be seen to widen even further when etched on top of the line grating (up to $\sim$ 60 nm on each side). Since the non-etched parts are smaller, the repetition length of the randomization unit is maintained.

The surface profile of the structure is shown in Figure 4. From several images similar to this, we conclude that the underlying periodic line grating had been etched 215 nm into the Si substrate, and that performing the extra etching of the randomization reticle only slightly influences the line grating heights. The etch depth of the random grating is approximately 150 nm.

B. Optical verification

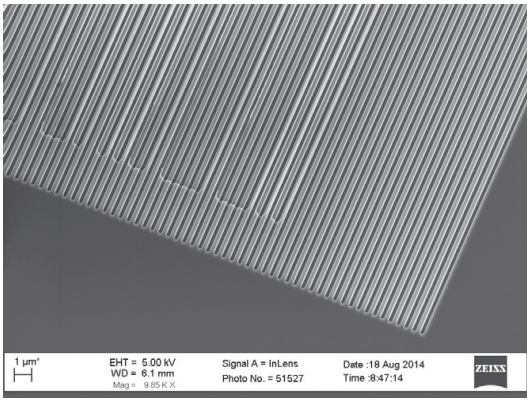
The sample reflectance was characterized by focusing an unpolarized, white light source through a fiber almost normal to its surface (only tilted enough to avoid incident light being reflected back into the source) and then scanning a 200 μm optical fiber in the plane of reflection, see Figure 5. The method used is described in detail elsewhere [21].

The small deviation from normal incidence means that the reflection plane due to invariance is slightly curved [22], but by visual inspection it was confirmed that the reflection plane was covered by the detecting fiber. It might have contributed to a decrease in intensity for larger angles. The result of the measurement is presented in Figure 6(e). The angular resolution when recording the measurements was 2 degrees, and the spectral resolution was better than 0.5 nm. For large angles, the received intensity was low, and the signal-to-noise ratio therefore worse than for small angles.

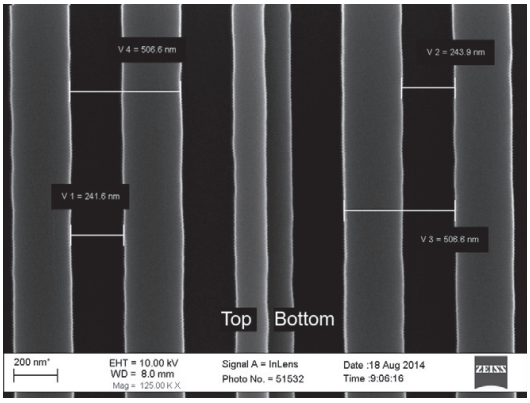
7. RESULTS ANALYSIS

Many of the observed features on the contour plots in Figure 6 can be described from the combination of the unit structure and the SAF approximation – that is, either they come from the choice of unit structure or choice of randomization. Figure 7 shows an annotated version of the FEM response and this can be used to provide further understanding on how to interpret the observed reflection data. First of all, it is noticed how the decrease in specular intensity around 450 nm is due to the choice of unit structure (see Figure 6(a) and 7), whereas the specular roll off at longer wavelength is caused by the SAF (see Figure 6(b)). In the first order grating mode, we see several intensity dips, which all can be ascribed either to unit structure or SAF (see Figure 7). Another set of distinct features in the numerical verification is the group of lines that correspond to higher order grating modes (multiples of 500 nm). These are caused by the randomization pattern (Figure 7 and 6(b)) and will be referred to as superperiod modes in the rest of this text. These superperiods are better quenched in the approximation than the verification. This is probably due to the periodic boundary conditions assumed when simulation one unit compared to the disturbed periodicity in the full simulation, making the destructive interference condition less ideal. A final important feature is the control of intensity distribution, which to a large extent is mediated by the unit cell structure (Figure 6(a) and 7). This arises because there is a weak wavelength dependence for the spread in intensity of the SAF at longer wavelengths (Figure 6(b).

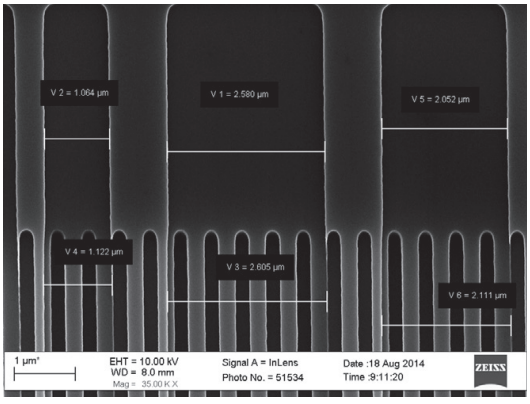
Comparing the measurement results in Figure 6(e) and Figure 6(c-d), we see that a fit between experimental data from the fabricated sample and reflection data arising from the FEM model is generally very good. We notice several features indicating this fit. Firstly, the grating modes from superperiods are still visible albeit at much lower intensity; this is probably due to the general smoothing effect observed in the measurements. Secondly, the spectral intensity profile of the first order grating mode is similar to that observed in the model. Some features



(a)



(b)



(c)

Fig. 3. SEM images of the produced structural color sample. (a) The structure overview shows the unit half-pitch grating with the random grating visible inside the half-pitch grating as undulations to the otherwise regular pattern. (b) Close-up of the unit grating structure at the edge of the random structure showing a step height change. The lateral dimensions of the unit grating does not change between steps. (c) Close-up of the random grating. The random grating is indeed random in its lateral dimension.

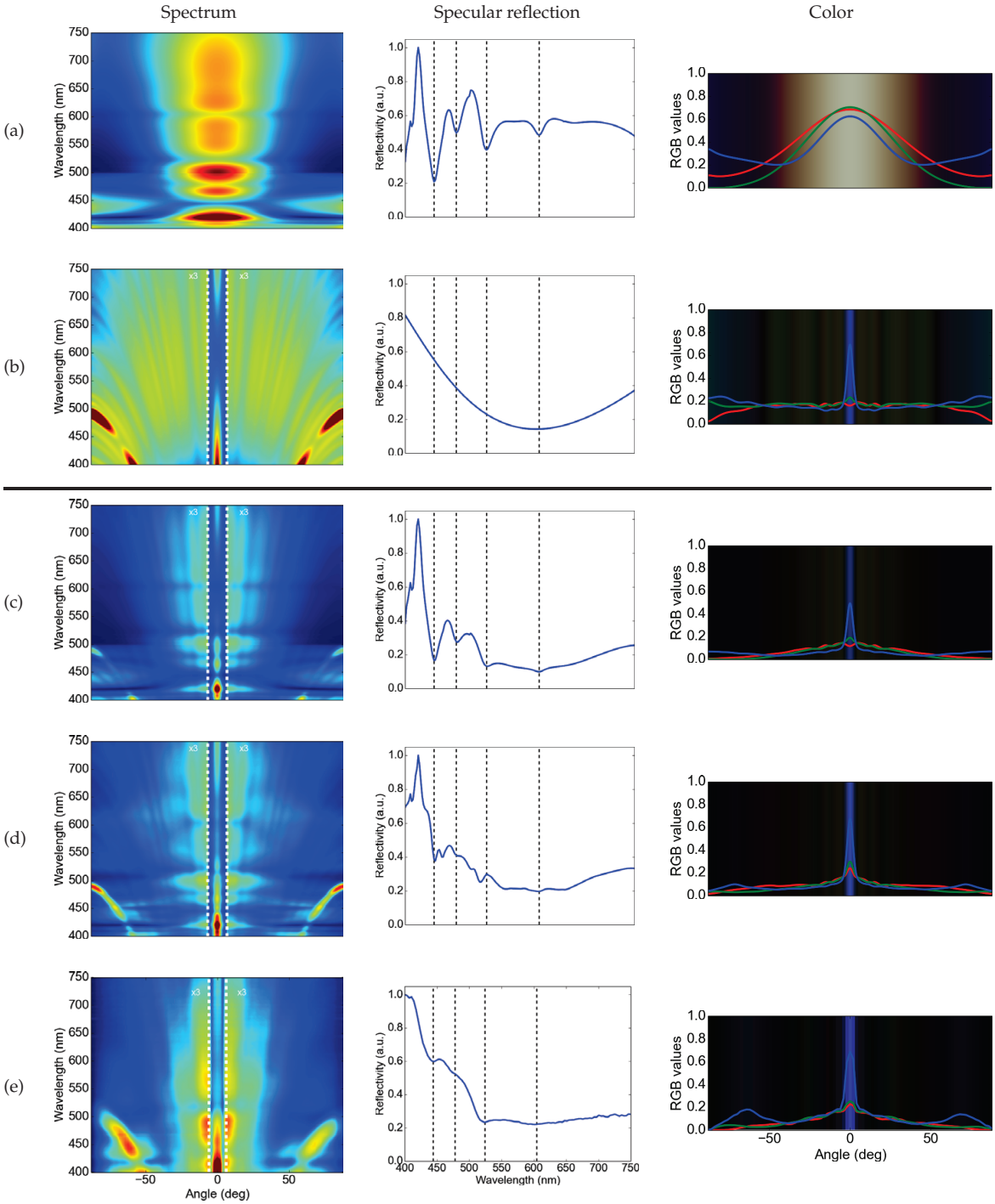


Fig. 6. (color online) Results of each design step in obtaining a structurally colored surface: (a) the response of the chosen unit structure; (b) the influence of randomization (SAF); (c) the product of these two then yields an approximate response; (d) a full wave FEM simulation verifies the result before production; (e) measurements on a produced sample is compared with the predicted result. All spectra are given in radiance (which here means intensity divided by cosine to the angle of reflection). All simulated data has been Gaussian blurred using an image filter as a simple way of incorporating the effect of the finite sized detector in the measurement setup. Note that some values in the contour plot are cropped in order to get a better color representation.

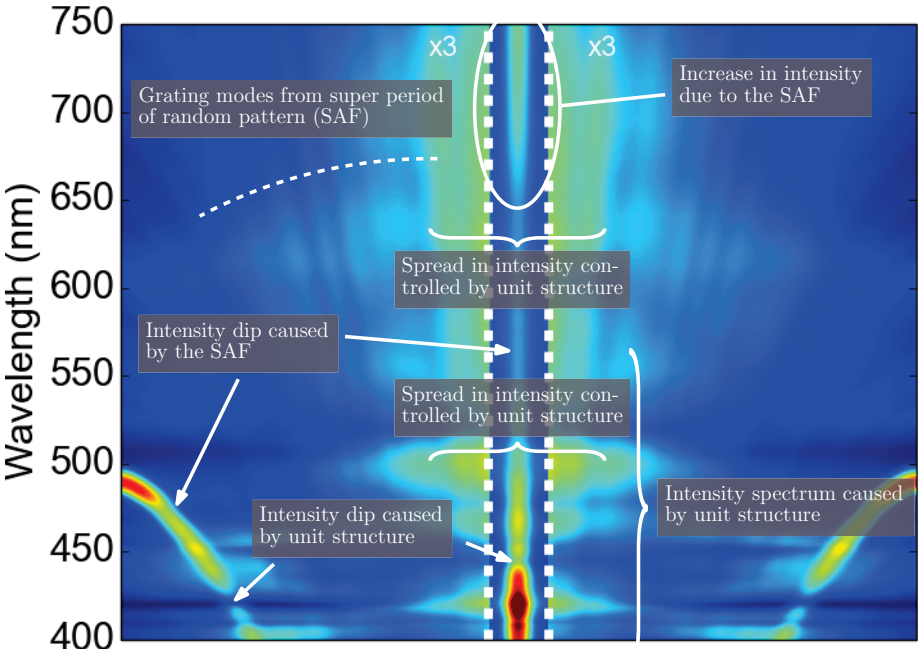


Fig. 7. (color online) An annotated version of the numerical verification presented in Figure 6(d) explaining where the effects visible on the plot originates from.

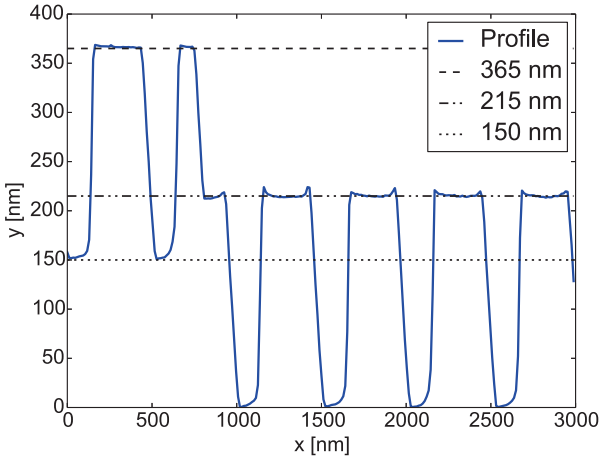


Fig. 4. Surface profile of the silicon structure extracted from AFM measurements.

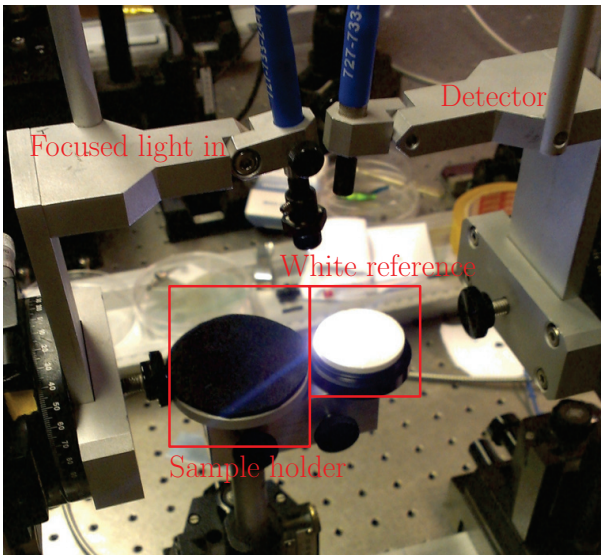


Fig. 5. (color online) Setup for reflectance characterization of the produced sample. The sample (or calibration reference) is lit by a focused white light source from a 1000 μm fiber and collected by a 200 μm fiber connected to a USB2000 Ocean Optics spectrometer.

indicating a clear difference between experimental data and the model are also present: the broad scattering around the specular direction for larger wavelengths is much less pronounced in the experimental data; the amount of scattered intensity around the specular direction for lower wavelengths is much more pronounced in the experimental data than in the model. Since the sample geometry is of high precision, we believe that the difference between model and measurements primarily is caused by nanometer-sized height variations in the sample. This decreases the flawless interference effects present in the simulation. This can also explain why, for example, the intensity minima in specular reflection are much less pronounced in the model compared to the measurements.

8. CONCLUSIONS

We have presented a method for fabricating surface structures that leads to a target structural color performance. The methods consist of several steps: a unit design phase; a randomization design phase; verification and production. This enables the fabrication of a structurally colored surface by the workflow shown in Figure 2. Despite some open questions, a general match between modeling and experiment was observed, verifying the overall design approach. In earlier work some of the design steps have been discussed in detail [15, 17, 19], and this should enable more advanced designs featuring more complex and appealing visual appearances. In the future these would be based on more complex structures – especially by extending the approach to 3D.

We believe that this approach makes it possible to explore the design space for different manufacturing technologies and materials. A long-term goal is to be able to create cheap, mass-manufacturable, environmentally friendly structurally colored surfaces based on polymer materials and mass-production techniques like hot embossing and injection molding. The low refractive index of these materials might limit total reflection compared to silicon but the technology is still relevant for applications like text, color symbols, security logos and decorative purposes.

FUNDING INFORMATION

The authors are grateful for the support from the Danish National Technology Foundation through the ODAAS and Nanoplast projects. Peter Vukusic was supported by USAF grant number FA9550-10-1-002.

ACKNOWLEDGEMENTS

The authors would like to thank Mark Heath and David Hudson of the Centre for Graphene Science, University of Exeter for help on sample preparation. Also Lesley Wears of College of Engineering, University of Exeter is thanked for help on characterization.

REFERENCES

1. M. Kolle, P. M. Salgard-Cunha, M. R. J. Scherer, F. Huang, P. Vukusic, S. Mahajan, J. J. Baumberg, and U. Steiner, "Mimicking the colourful wing scale structure of the *Papilio blumei* butterfly," Nature Nanotechnology **5**, 511–15 (2010).
2. C. I. Aguirre, E. Reguera, and A. Stein, "Colloidal photonic crystal pigments with low angle dependence," Applied Materials & Interfaces **2**, 3257–62 (2010).
3. C. E. Finlayson and J. J. Baumberg, "Polymer opals as novel photonic materials," Society of Chemical Industry **62**, 1403–07 (2013).
4. N. Koay, I. B. Burgess, T. M. Kay, B. A. Nerger, M. Miles-Rossouw, T. Shirman, T. L. Vu, G. England, K. R. Phillips, S. Utech, N. Vogel, M. Kolle, and J. Aizenberg, "Hierarchical structural control of visual properties in self-assembled photonic-plasmonic pigments," Optics Express **22**, 27750–68 (2014).
5. A. Saito, Y. Miyamura, Y. Ishikawa, and J. M. andMegumi Akai-Kasaya andYuji Kuwahara, "Reproduction, mass-production, and control of the *Morpho*-butterfly's blue," in "Advanced Fabrication Technologies for Micro/Nano Optics and Photonics II," , vol. 7205 (2009), vol. 7205, pp. 720506–9.
6. J. Yip, S.-P. Ng, and K.-H. Wong, "Brilliant whiteness surfaces from electrospun nanofiber webs," Textile Research Journal **79**, 771–79 (2009).
7. L. Yisen, C. Yi, L. Zhiyuan, H. Xing, and L. Yi, "Structural coloring of aluminum," Electrochemistry Communications **13**, 1336–39 (2011).
8. M. A. Kats, R. Blanchard, P. Genevet, and F. Capasso, "Nanometre optical coatings based on strong interference effects in highly absorbing media," Nature Materials **12**, 20–24 (2012).
9. M. Aryal, D.-H. Ko, J. R. Tumbleston, A. Gadisa, E. T. Samulski, and R. Lopez, "Large area nanofabrication of butterfly wing's three dimensional ultrastructures," Journal of Vacuum Science & Technology B **30**, 1–7 (2012).
10. K. Kumar, H. Duan, R. S. Hedge, S. C. W. Koh, J. N. Wei, and J. K. W. Yang, "Printing colour at the optical diffraction limit," Nature Nanotechnology **7**, 557–61 (2012).
11. M. Kolle, A. Lethbridge, M. Kreysing, J. J. Baumberg, J. Aizenberg, and P. Vukusic, "Bio-inspired band-gap tunable elastic optical multilayer fibers," Advanced Materials (Weinheim, Germany) **25**, 2239–45 (2013).
12. Y. Chuo, C. Landrock, B. Omrane, D. Hohertz, S. V. Grayli, K. Kavanagh, and B. Kaminska, "Rapid fabrication of nanostructured quartz stamps," Nanotechnology **24**, 1–10 (2013).
13. J. S. Clausen, E. Højlund-Nielsen, A. B. Christiansen, S. Yazdi, M. Grajower, H. Taha, U. Levy, A. Kristensen, and N. A. Mortensen, "Plasmonic metasurfaces for coloration of plastic consumer products," Nano Letters **14**, 4499–504 (2014).
14. N. Schneider, C. Zeiger, A. Kolew, M. Schneider, J. Leuthold, H. Hölscher, and M. Worgull, "Nanothermoforming of hierarchical optical components utilizing shape memory polymers as active molds," Optical Materials Express **4**, 1895–902 (2014).
15. V. E. Johansen, "Optical role of randomness for structured surfaces," Applied Optics **53**, 2405–15 (2014).
16. A. Saito, M. Yonezawa, J. Murase, S. Juodkazis, V. Mizeikis, M. Akai-Kasaya, and Y. Kuwahara, "Numerical analysis on the optical role of nano-randomness on the *Morpho* butterfly's scale," Journal of Nanoscience and Nanotechnology **11**, 2785–92 (2011).
17. J. Andkjær, V. E. Johansen, K. S. Friis, and O. Sigmund, "Inverse design of nanostructured surfaces for color effects," Journal of the Optical Society of America B **31**, 164–74 (2014).
18. K. S. Friis and O. Sigmund, "Robust topology design of periodic grating surfaces," Journal of the Optical Society of America B **29**, 2935–43 (2012).
19. V. E. Johansen, J. Andkjær, and O. Sigmund, "Design of structurally colored surfaces based on scalar diffraction

theory," Journal of the Optical Society of America B **31**, 207–17 (2014).
20. E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, 1997).
21. P. Vukusic and D. G. Stavenga, "Physical methods for investigating structural colours in biological systems," Journal of the Royal Society Interface **6**, S133–48 (2009).
22. J. E. Harvey and C. L. Vernold, "Transfer function characterization of scattering surfaces: Revisited," in "SPIE," , vol. 3141 (1997), vol. 3141, pp. 113–27.

Publication P5

Preparing the Generalized Harvey-Shack rough surface scattering method for use with the Discrete Ordinates Method

Preparing the generalized Harvey–Shack rough surface scattering method for use with the discrete ordinates method

Villads Egede Johansen

Technical University of Denmark, Department of Mechanical Engineering, 2800 Kgs. Lyngby, Denmark (vejo@mek.dtu.dk)

Received September 5, 2014; revised November 4, 2014; accepted November 19, 2014; posted November 20, 2014 (Doc. ID 221781); published January 8, 2015

The paper shows how to implement the generalized Harvey–Shack (GHS) method for isotropic rough surfaces discretized in a polar coordinate system and approximated using Fourier series. This is particularly relevant for the use of the GHS method as a boundary condition for radiative transfer problems in slab geometries, where the discrete ordinates method can be applied to solve the problem. Furthermore, such an implementation is a more convenient discretization of the problem than the traditional direction cosine space that has its strengths in analytical problems and intuitive understanding (mainly due to its translation invariance). A computer implementation of scattering from a Gaussian rough surface with Gaussian autocovariance written in Python is included at the end of the paper.

OCIS codes: (290.5835) Scattering, Harvey; (240.5770) Roughness; (290.1483) BSDF, BRDF, and BTDF; (010.5620) Radiative transfer; (110.7050) Turbid media.

http://dx.doi.org/10.1364/JOSAA.32.000186

1. INTRODUCTION

Scattering from rough surfaces continues to be an active area of research in optics [1]. A multitude of approaches to solve roughness scattering problems exists [2], all with their advantages and disadvantages—most often tailored to a range of applications. One set of often used methods is based on single scattering theories [1], which historically have been popular due to their simplicity. However, there is a natural limit to which effects these theories can capture since multiple scattering will occur when surface perturbations get more prominent, and the famous Rayleigh criterion is no longer fulfilled.

The best approach to overcome these issues and still keep a simple formulation seems to be the recent generalized Harvey–Shack (GHS) method [3], which is based on the principle of nonparaxial scalar diffraction theory [4], which significantly improves scattering predictions for large angle scattering. Harvey has on several occasions [3–7] indicated its superior predictions compared to paraxial scalar diffraction theory and older single scattering rough surface theories. This also includes results showing good fit with experimental data for moderately rough surfaces.

When the rough surface calculation in itself is a part of a larger model [8–12] and results are needed within a reasonable time, GHS is therefore a good approach due to its accuracy and low computational cost compared to simple scalar methods such as Rayleigh–Rice (low accuracy, no energy conservation) and full wave methods (high computational cost). A sketch of such a model is shown in Fig. 1(a). In the paper it will be shown how GHS can be implemented for use in the radiative transfer model when solved using the so-called discrete ordinates method (DOM) [9,13,14]. This enables GHS to be used efficiently when, for example, determining reflection from turbid media in slab geometries with rough surfaces. Furthermore, this approach seems more suitable for implementing the GHS method in general.

A secondary purpose of the paper is to present a minimalistic, ready-to-use implementation of the GHS method, in order to make it accessible for researchers in need of a scattering model for rough surfaces. An implementation of the GHS method is therefore given in Appendix A.

The rest of the paper is organized as follows. Section 2 introduces notation that will be used subsequently, Section 3 gives a short description of the DOM and radiative transfer, Section 4 introduces the necessary knowledge on rough surfaces needed for the GHS method, and Section 5 states all the relevant equations needed to implement GHS. After these recapitulations follows Section 6, which reformulates GHS in polar coordinates and Fourier series, thus having prepared it for DOM; Section 7, which explains and comments on the Python implementation of the GHS method presented in Appendix A; and Section 8, which concludes on the reformulation and implementation.

2. NOTATION

To be consistent with similar writings [6], all variables subscripted i refer to the incident angle, a subscript of o refers to the specular reflection angle (meaning $\theta_i = -\theta_o$), and s refers to scattered angles. Furthermore, a hatted variable means that it is scaled with respect to the electromagnetic wavelength of interest, λ. For example, $\hat{x} = x/\lambda$, $\hat{\sigma} = \sigma/\lambda$. The Fourier transform operator $\mathcal{F}$ is defined as

$$\mathcal{F}\{H(\hat{x},\hat{y})\}(\alpha,\beta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(\hat{x},\hat{y})e^{-2\pi i(\hat{x}\alpha+\hat{y}\beta)}\mathrm{d}\hat{x}\mathrm{d}\hat{y}. \quad (1)$$

1084-7529/15/020186-09$15.00/0

In this text $\hat{x}$, $\hat{y}$ are always the nontransformed variables. In many cases some variables are uniquely defined by others, and therefore not all variables are written as independent variables of a function. This is, for example, the case where $\gamma_i = -\gamma_o$, $\gamma = \cos\theta$ or the condition in Eq. (12).

3. RADIATIVE TRANSFER FOR SLAB GEOMETRIES

The source free, isotropic radiative transfer equation (RTE) for a slab geometry as seen in Fig. 1(a) describes the angular redistribution/transfer of radiance u with respect to the optical depth $\tau \propto z$ in a turbid/scattering medium described by the phase function p such that

$$-\mu\frac{\mathrm{d}u(\tau,\mu,\phi)}{\mathrm{d}\tau} = -u(\tau,\mu,\phi) + \frac{\Omega}{4\pi}\int_{-1}^{1}\int_{0}^{2\pi} p(\mu',\phi',\mu,\phi)u(\tau,\mu',\phi')\mu'\mathrm{d}\mu'\mathrm{d}\phi', \tag{2}$$

where $\mu = \cos\theta$, ϕ is as defined in Fig. 1 and Ω is the so-called albedo. This equation can, for example, be used to describe light scattering of paint or skin. To cover the effect of the rough boundaries in Fig. 1(a), the boundary conditions for the RTE are normally modified to take this into account [9,17,18]. Whereas rough surfaces for, e.g., ocean analysis have received much attention [17,18], it is hard to find implementations for optically rough surfaces [9]. The author has not been able to find any implementation that can be based on actual roughness measurements and that preserves a collimated beam component in the case of low roughness. The GHS method satisfies both criteria, and can therefore be useful as a rough boundary condition for these cases.

A. Discrete Ordinates Method

The DOM was introduced by Chandrasekhar [13,19] to solve the RTE for plane-parallel/slab geometries. The method has later been modified for efficient, stable numerical implementations [14,20] and also applied to problems of higher dimensionality [19]. DOM discretizes μ in its Gauss points (ordinates) to facilitate numerical integration [13]. By defining the Gauss points x_i and their corresponding weights w_i, integrals can be approximated by Gaussian quadrature:

$$\int_{-1}^{1} f(x) \approx \sum_{i=1}^{N} w_i f(x_i). \tag{3}$$

Most modern math-oriented programming languages contain methods for calculating Gauss points and their corresponding weights. The integration rule can of course be scaled. At the same time, the radiance is linearized using a Fourier series expansion in the ϕ direction to express the solution. For a 2π periodic function f that can be expressed as a Fourier series it holds that

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx), \tag{4}$$

where

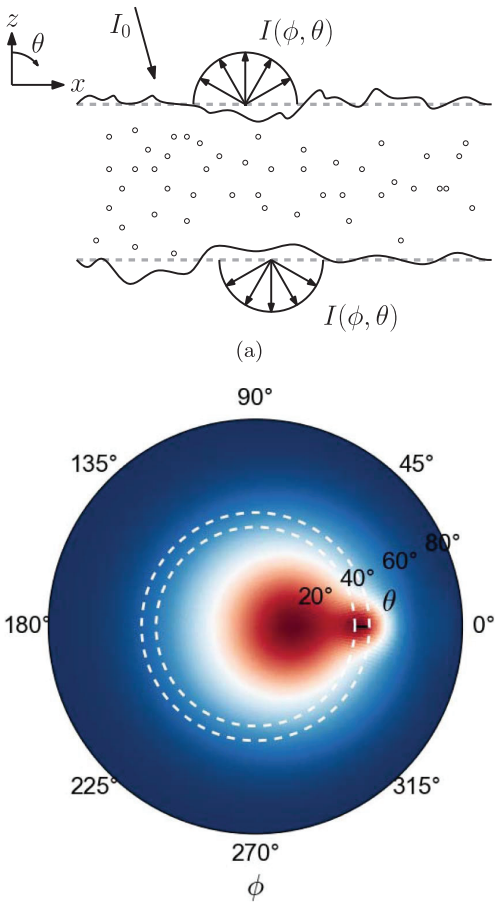


Fig. 1. (a) Example of a slab geometry with rough surfaces and a scattering/turbid medium in between (here depicted by particles). This model is, for example, used to represent human tissue [9], the ocean [15], and paint [16]. (b) Example of reflected intensity from such a system. Under the assumption that the rough surfaces as well as the turbid medium scatter isotropically, the corresponding reflection will be symmetric in the ϕ direction. This is shown here by dashed lines indicating constant ϕ circles as well as a black bar showing the symmetry line.

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx, \mathrm{d}x \qquad n = 0,1,2,\ldots, \tag{5}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx, \mathrm{d}x \qquad n = 1,2,3,\ldots. \tag{6}$$

In order to obtain a finite linear system, the summation is truncated at a certain order N. This is justified since, if the absolute squared integral of the derivative of f is finite, the error rolls off with at least a factor of $1/\sqrt{N}$ [21, p. 127].

For isotropic cases, the solution is symmetric in ϕ with respect to the plane of incidence, which means that $b_n = 0\,\forall n \in \mathbb{N}$. The symmetry is sketched in Fig. 1(b). To be able to impose the rough surface boundary conditions, it is necessary to discretize the GHS method in the same way.

4. ISOTROPIC ROUGH SURFACES

All rough surfaces described by the GHS method are assumed to be isotropic and to have a Gaussian surface height distribution (with a zero mean height) [6]. This means that a root-mean-square (rms) surface roughness σ_s can be associated with the roughness properties of the surface, since it is then the standard deviation of the height distribution. Furthermore, a so-called surface autocovariance (ACV) function can be ascribed to the surface [6]. The ACV does not need to have a Gaussian distribution. Since not all features contribute to light scattering (such as scratches or other surface "finish" effects), the ACV is also used to calculate the so-called relevant roughness as described below, which corrects for this effect.

A. Relative Roughness

Considering the surface power spectral density (PSD) function, which is the Fourier transform of the ACV, it is found that some of these spatial frequencies will only contribute to evanescent modes [6], which do not contain any energy. This happens when the spatial frequencies are greater than $1/\lambda$ [6], and for a given incident angle θ_i, the relevant rms surface roughness is therefore limited to

$$\sigma_{\mathrm{rel}} = \sqrt{\int_{-1/\lambda+f_o}^{1/\lambda+f_o} \int_{-\sqrt{1/\lambda^2-(f_x-f_o)^2}}^{\sqrt{1/\lambda^2-(f_x-f_o)^2}} \mathrm{PSD}(f_x,f_y)\mathrm{d}f_y\mathrm{d}f_x}, \tag{7}$$

where

$$f_o = \frac{-\sin\theta_i}{\lambda}. \tag{8}$$

The above expressions can be understood as integrating the PSD in a disc with radius $1/\lambda$ centered at f_o instead of integrating over the whole f_x, f_y plane. This ensures that only roughness that contributes to scattering is considered. A more elaborate description is found in [3,6].

B. Rough Surface with Gaussian ACV

As a verification example for the implementation, a rough surface described by a Gaussian ACV is used so that the results can be compared against a reference [6]. The ACV is given by

$$C_s(\hat{x},\hat{y}) = \hat{\sigma}_s^2 e^{-\hat{r}^2/\ell_c^2}, \tag{9}$$

where $\hat{r} = \sqrt{\hat{x}^2+\hat{y}^2}$ and ℓ_c is the correlation length of the surface [6]. The Fourier transform of this circular symmetric function can be looked up in most books or Web pages treating two-dimensional Fourier transforms, and the result is that

$$\mathrm{PSD}(f_r) = \pi\ell_c^2\sigma_s^2 e^{-(\pi\ell_c f_r)^2}, \tag{10}$$

where $f_r = \sqrt{f_x^2+f_y^2}$.

5. OVERVIEW OF THE GHS METHOD

Before deriving the GHS method for polar coordinates, the relevant equations for calculating light reflection from isotropic rough surfaces will be stated [6]. A thorough treatment of GHS for rough surfaces is found in [3,6] alongside a few validation examples on experimental data. A numerical validation of the method is available in [7].

Defining a traditional spherical coordinate system as sketched in Fig. 2, where $\phi \in [0, 2\pi]$, $\theta \in [0, \pi]$, the so-called direction cosines can be defined as

$$\alpha = \sin\theta\cos\phi, \qquad \beta = \sin\theta\sin\phi, \qquad \gamma = \cos\theta. \tag{11}$$

Note that this description is overdetermined since three variables are used to describe two variables—for example, this implies that

$$\gamma = \mathrm{sign}(\pi/2-\theta)\sqrt{1-\alpha^2+\beta^2}, \tag{12}$$

where $\mathrm{sign}(x)$ is 1 for $x \geq 0$ and -1 otherwise.

Having defined the coordinate system, it is now possible to describe the reflected or transmitted intensity using GHS. For simplicity only reflection will be considered here, and in [3,6] the few considerations needed for calculating transmittance are described. The reflected intensity is given as [6]

$$I(\theta_i,\phi_i,\theta_s,\phi_s,\lambda) = \cos(\theta_s)\mathrm{BRDF}(\theta_i,\phi_i,\theta_s,\phi_s,\lambda), \tag{13}$$

where BRDF is the so-called bidirectional reflection distribution function (BRDF), which describes the distribution of reflected radiance for light incident at (θ_i,ϕ_i). The BRDF generated by the rough surface is then described by [6]

$$\mathrm{BRDF}(\theta_i,\phi_i,\theta_s,\phi_s,\lambda) = R(\theta_i)\mathrm{ASF}(\theta_i,\phi_i,\theta_s,\phi_s)|_\lambda, \tag{14}$$

where R is the Fresnel reflection calculated for a given incident angle, and ASF is the so-called angle spread function. The ASF can be represented using direction cosines, $\mathrm{ASF}(\alpha_i,\beta_i,\alpha_s,\beta_s)$. Without losing generality $\alpha_i = 0$ is assumed. This is legal due to the isotropic surface assumption that makes the reflection from the rough surface indifferent to rotation in the ϕ direction. The ASF is then defined as [6]

$$\mathrm{ASF}(\beta_i,\alpha_s,\beta_s) = A(\gamma_i,\gamma_o)\delta(\alpha_s,\beta_s-\beta_o) + K(\gamma_i)S(\beta_i,\alpha_s,\beta_s). \tag{15}$$

Here, A describes the energy left in the specular reflection and is given as

$$A(\gamma_i,\gamma_s) = e^{-[2\pi(n_1\gamma_i-n_2\gamma_s)\hat{\sigma}_{\mathrm{rel}}]^2}, \tag{16}$$

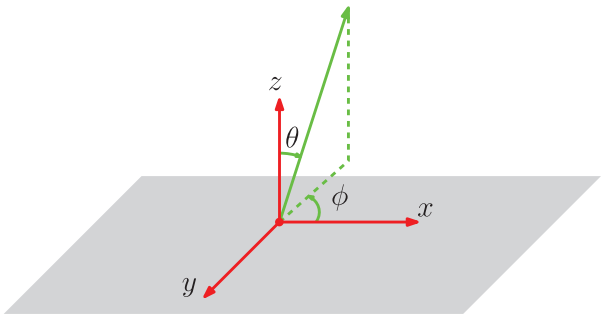


Fig. 2. Coordinate system used in this text.

where $\hat{\sigma}_{\text{rel}} = \sigma_{\text{rel}}/\lambda$ was calculated in the unscaled version in Section 4.A and n_1, n_2 are the refractive indices of the material from the background medium and the rough surface material, respectively. For reflection $n_2 := -n_1$. S describes the distribution of the scattered radiance and is given as

$$S(\beta_i, \alpha_s, \beta_s) = B(\gamma_i, \gamma_s)\mathcal{F}\{G(\hat{x}, \hat{y}, \gamma_i, \gamma_s)e^{-i2\pi\beta_o\hat{y}}\}, \tag{17}$$

where

$$B(\gamma_i, \gamma_s) = 1 - A(\gamma_i, \gamma_s) \tag{18}$$

and

$$G(\hat{x}, \hat{y}, \gamma_i, \gamma_s) = \frac{e^{[2\pi(n_1\gamma_i - n_2\gamma_s)\sigma_{\text{rel}}/\sigma_s]^2 C_s(\hat{x},\hat{y})} - 1}{e^{[2\pi(n_1\gamma_i - n_2\gamma_s)\hat{\sigma}_{\text{rel}}]^2} - 1}, \tag{19}$$

where C_s is the ACV function of the rough surface. A concrete choice of ACV for the later implementation was introduced in Section 4.B.

K in Eq. (15) is used to normalize S by assuming no energy is lost in evanescent modes, and also making sure that energy not left in the specular reflection is in the scattered term, and is therefore defined as

$$K(\gamma_i) = B(\gamma_i, \gamma_o)\left(\int_{-1}^{1}\int_{-\sqrt{1-\alpha^2}}^{\sqrt{1-\alpha^2}} S(\beta_i, \alpha, \beta)\mathrm{d}\alpha\mathrm{d}\beta\right)^{-1}, \tag{20}$$

where the integral is equivalent to integrating over the hemisphere $\phi \in [0, 2\pi]$, $\theta \in [0, \pi/2]$.

A. Calculating S Using Translated Hankel Transforms

For normal incidence, $\beta_i = 0$, S is a circular symmetric (that is, constant w.r.t. ϕ) function due to the isotropy assumption. This means that the Fourier transform in Eq. (17) can be calculated efficiently by applying the Hankel transform of order zero, and this feature is preferable to retain. Therefore, S_{rad} is defined such that

$$S_{\text{rad}}(\beta_i, \alpha_s, \beta_s) = B(\gamma_i, \gamma_s)\mathcal{F}\{G(\hat{x}, \hat{y}, \gamma_i, \gamma_s)\}, \tag{21}$$

which can then be calculated efficiently using a Hankel transform. Details are shown in the implementation. Furthermore, the modulation (exponential function) in Eq. (17) corresponds to a translation in the transformed domain, which means S can be expressed as

$$S(\beta_i, \alpha_s, \beta_s) = S_{\text{rad}}(\beta_i, \alpha_s, \beta_s - \beta_o). \tag{22}$$

This will later on make it easy to calculate S, which is best seen in the actual implementation in Appendix A. By this we have a complete and efficient description of the BRDF of the rough surface. The task is then to convert the method such that it is represented by its Fourier components in a μ, ϕ coordinate system suitable for DOM.

6. REFORMULATING GHS FOR DOM

Reformulating the GHS method in a μ, ϕ coordinate system gives cumbersome expressions, and a $\rho = \sin\theta, \phi$ system is preferred. This also places $\theta = 0°$ at the origin, which is more intuitive. In a discretized version, this is not a problem, as long as all points ρ_i are chosen to match a Gauss point μ_i. Integrals still have to be reformulated to fit the Gauss points of μ, which will also be shown in the following. The extra step is needed to comply with DOM, but *if the calculations are not to be used with DOM, the author recommends to discretize ρ in its Gauss points instead to get a simpler implementation.*

A. Updating Coordinates and Translation

In this section the problem is reformulated in terms of ρ, θ. To begin with, note that

$$\beta_o = \sin\theta_o = \rho_o, \quad \text{for } \phi = \pi/2, \tag{23}$$

so that β_o translation can be directly interchanged with ρ_o translation along the $\phi = \pi/2$ line. (This would not have been possible if choosing μ.) A Cartesian translation of the radial component in a polar coordinate system is done by updating the radial component to

$$\rho' = \sqrt{\rho^2 + \rho_o^2 - 2\rho\rho_o\cos\phi}, \tag{24}$$

where ρ_o is the length of the translation and ρ' is the updated variable. This transformation also involves ϕ, since such a translation naturally breaks the circular symmetry (for all but the constant case).

The relation between the direction cosines and ρ, ϕ is given by

$$\alpha = \rho\cos\phi, \qquad \beta = \rho\sin\phi, \qquad \gamma = \text{sign}(\rho)\sqrt{1-\rho^2}. \tag{25}$$

If the changes are introduced in Eq. (21), then it is easy to calculate

$$S_r(\rho_i, \rho_s) := S_{\text{rad}}(\beta_i = \rho_i, \alpha_s = 0, \beta_s = \rho_s). \tag{26}$$

Note that there is no dependence of ϕ_s for S_r due to the isotropic scattering, and it is therefore chosen to put $\phi_s = \pi/2$. This then means that S can be expressed as

$$S(\rho_i, \rho_s, \phi_s) = S_r\left(\rho_i, \sqrt{\rho_s^2 + \rho_o^2 - 2\rho_s\rho_o\cos\phi_s}\right). \tag{27}$$

B. Fourier Series Expansion

Due to DOM, the solution has to be expanded into its Fourier series components. From Section 3 it is found that the b_n's all are zero. Expanding the Fourier series, it can be written as

$$S(\rho_i, \rho_s, \phi_s) = \frac{1}{2}s_0(\rho_i, \rho_s) + \sum_{k=1}^{\infty} s_k(\rho_i, \rho_s)\cos k\phi_s, \tag{28}$$

where

$$s_k(\rho_i, \rho_s) = \frac{2}{\pi}\int_0^{\pi} S(\rho_i, \rho_s, \phi)\cos k\phi\mathrm{d}\phi, \tag{29}$$

where the integral will be solved numerically by sampling ϕ at Gauss points and then using the Gaussian quadrature rule as described in Section 3.

C. Calculating K

For calculating K, the integral defined in Eq. (20) now needs to be converted to the same coordinate system. The Jacobian associated with shifting to polar coordinates is given by

$$J = \frac{\partial\alpha}{\partial\rho}\frac{\partial\beta}{\partial\phi} - \frac{\partial\alpha}{\partial\phi}\frac{\partial\beta}{\partial\rho} = \phi\,\sin\theta\cos^2\phi + \sin\theta\sin^2\phi = \sin\theta = \rho, \tag{30}$$

which means that K can be calculated as

$$\begin{aligned} K(\gamma_i) &= B(\gamma_i,\gamma_o)\left(\int_0^1\int_0^{2\pi} S(\rho_i,\rho,\phi_s)\rho\mathrm{d}\phi\mathrm{d}\rho\right)^{-1} \\ &= B(\gamma_i,\gamma_o)\left(\pi\int_0^1 s_0(\rho_i,\rho)\rho\mathrm{d}\rho\right)^{-1}. \end{aligned} \tag{31}$$

The last equation is obtained since

$$\int_0^{2\pi} s_k(\rho_i,\rho_s)\cos k\phi_s = 0 \quad \text{for } k = 1,2,\ldots. \tag{32}$$

As is known from DOM, information on the energy of S is stored in the zero-order term, whereas all higher-order terms are only used to determine the shape of the reflection.

D. Calculating K by μ Integration

For the DOM purpose, it is necessary to integrate over μ instead of ρ. Remembering that $\rho = \sqrt{1-\mu^2}$ the Jacobian for this coordinate shift is

$$J = \frac{\partial\rho}{\partial\mu} = -\frac{\mu}{\sqrt{1-\mu^2}}. \tag{33}$$

This means that Eq. (31) can be expressed as

$$\begin{aligned} K(\gamma_i) &= B(\gamma_i,\gamma_o)\left(\pi\int_0^1 s_0\left(\rho_i,\sqrt{1-\mu^2}\right)\rho\frac{\mu}{\sqrt{1-\mu^2}}\mathrm{d}\mu\right)^{-1} \\ &= B(\gamma_i,\gamma_o)\left(\pi\int_0^1 s_0\left(\rho_i,\sqrt{1-\mu^2}\right)\mu\mathrm{d}\mu\right)^{-1}. \end{aligned} \tag{34}$$

E. Calculating Relative Roughness in Polar Coordinates

Even though this is not strictly necessary, it is convenient to convert the integral of σ_{rel} in Eq. (7) to an integral in polar coordinates, since the same principle then can be applied. The integral in Eq. (7) corresponds to the integral of a disc with radius $1/\lambda$ of the PSD translated by f_o, and can therefore be calculated—using the same principle as when calculating K—as

$$\sigma_{\mathrm{rel}} = \sqrt{\int_0^{1/\lambda}\int_0^{2\pi} \mathrm{PSD}_r(f_r,\phi_r,f_o) f_r \mathrm{d}\phi_r \mathrm{d}f_r}, \tag{35}$$

where

$$\mathrm{PSD}_r(f_r,\phi_r,f_o) = \mathrm{PSD}\left(\sqrt{f_r^2 + f_o^2 - 2f_r f_o\,\cos(\phi_r)}\right). \tag{36}$$

This finishes all reformulation needed for implementation, and the s_k's that are necessary coefficients for the DOM are now found. To ease implementation for others and to test the above reformulation, a numerical implementation is presented in the following.

7. IMPLEMENTATION

In Appendix A, a Python implementation of the above method is given. To run the program, save the code in Appendixes A.1 and A.2 as, respectively, `cosineHarveyShack.py` and `HankelTransformSimple.py` in the same folder, and run `cosineHarveyShack.py` using Python. In Linux the command would be

`python cosineHarveyShack.py`, and then one of the results shown in Fig. 3(a) will appear (with less formating). The code requires `numpy` and `matplotlib` to be installed. Older versions of `numpy` do not contain the function `leggauss` to find Gauss points and weights; in this case the code tries to fall back on a `scipy` implementation of the same function, which requires that `scipy` is also installed. The code has been tested under Ubuntu Linux 14.04 64-bit in Python versions 2.7.6 and 3.4.0, and Windows 7 64-bit using Anaconda 2.0.1. Since Anaconda and other Python implementations are available for OS X as well, the software is expected to run on Mac.

The code is written to be compact, straightforward to read, and ready to use and extend. This means that a few performance tweaks have been left out due to simplicity, where the most notable is the lack of Fresnel reflection and efficient calculation of the Hankel transform, which will be explained in Section 7.C. In the following sections, the implementation of the code is explained.

A. Note on `Numpy` Matrix Multiplication for MATLAB Users

For MATLAB users inexperienced with `numpy` for Python, it is worth noticing that Python array/matrix indexing starts at 0 instead of 1. Furthermore there is no distinct * and .* operator. Instead, * indicates element-wise multiplication if the operands are arrays and matrix multiplication if the operands are matrices. In case the arrays have different (but compatible) shapes, the result is a higher-order array. Since all `numpy` data structures in this code are initialized as arrays, * and / are always performed element-wise.

B. Main Program Explained

line 1–10 Import of modules: `numpy` for calculation, `matplotlib` for visualization. A function, `leggauss` to calculate Gauss points and weights from either `numpy` or `scipy`. Note that `HankelLibSimple` is the custom module in Appendix A.2.

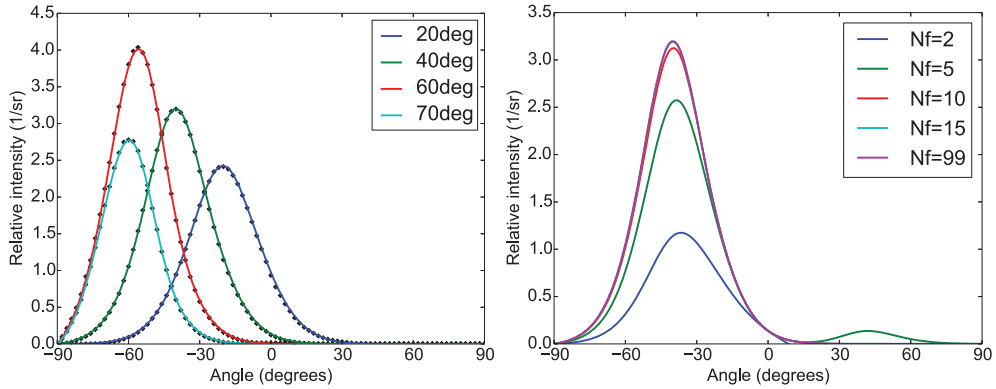
line 11–23 `main()` is the main routine that is run when the program is run; see lines 112–113. The input parameters are defined here:

`lam`	λ, wavelength;
`sigmas`	σ_s, rms roughness;
`lc`	ℓ_c, roughness correlation length;
`thetai`	θ_i, incident angle;
`n1`	n_1, refractive index for material 1;
`N`	N, number of Gauss points for μ;
`Nf`	N_f, number of Fourier expansion terms.

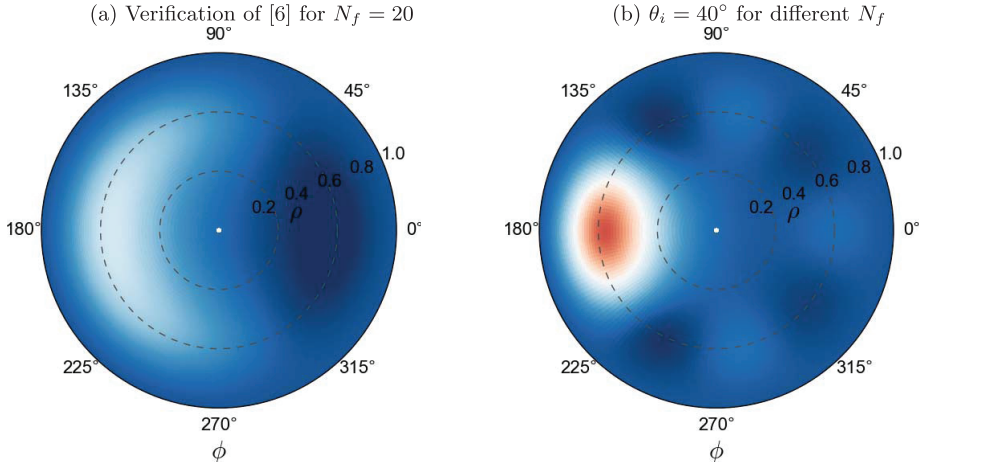
line 24–30 Discretization of the coordinate system in Gauss points. `phi` is transposed such that when multiplied by `rho` or `mu` a matrix multiplication will be returned.

line 31–36 Calculation of σ_{rel} based on Eq. (7)—using the Gauss quadrature rule for integration.

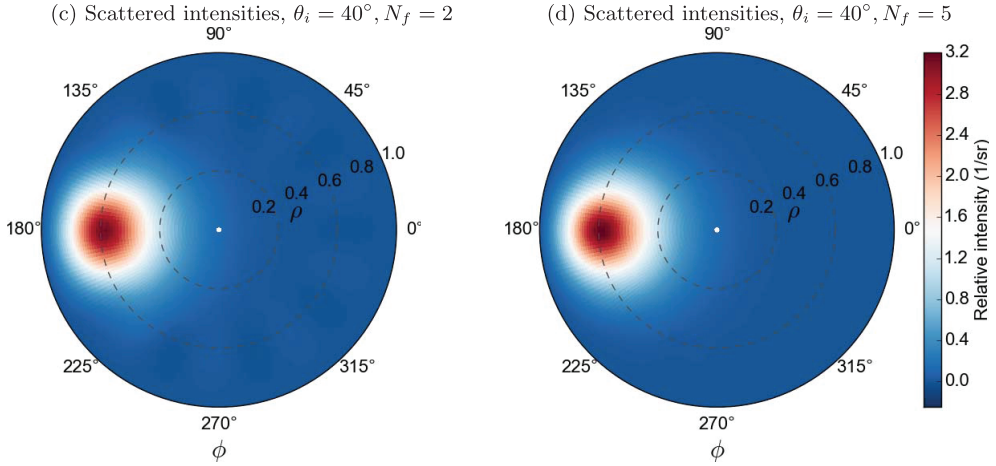
line 37–47 Precalculation and initialization of variables for later use. Note in particular that `cosk` is a matrix containing



(a) Verification of [6] for $N_f = 20$ (b) $\theta_i = 40°$ for different N_f



(c) Scattered intensities, $\theta_i = 40°, N_f = 2$ (d) Scattered intensities, $\theta_i = 40°, N_f = 5$



(e) Scattered intensities, $\theta_i = 40°, N_f = 10$ (f) Scattered intensities, $\theta_i = 40°, N_f = 15$

Fig. 3. Results calculated from the code in Appendix A—all settings are the same except for θ_i and N_f. (a) verifies the implementation by reproducing the results from [6], which is indicated by black diamonds. Mismatch is due to error in the digitizing process. (b) shows the results of calculating the same example for different numbers of cosine terms, N_f. (c)–(f) show the reflected intensities for the whole ρ, ϕ range.

$\cos(k\phi)$ for all discretized points of k and ϕ. This is used for the Fourier expansion calculations.

line 48–51 Setup of the Hankel transform class for repeated use in the following loop.

line 52–61 First part of loop over ρ_s, since Eqs. (21) and (27) dictate a new Hankel transform for every value of ρ_s (but not ϕ_s). Note how `Cs` in line 58 is evaluated using the radial coordinate system defined by the `HankelTransform` class. `fht` is short for *forward Hankel transform*.

line 62–68 Last part of loop over ρ_s. Here `Srad` (S_r) is interpolated onto the relevant part of `rhom` (ρ'). Since the interpolation requires monotone, increasing values, the

coordinates are first sorted, then interpolated, and then sorted back. In line 67 `sk` is calculated from Eq. (29) by the Gauss quadrature rule.

line 69–72 Calculation of K using Eq. (20), and B using Eq. (18).

line 73–86 Evaluation of S in the plane of incident light using Eq. (28) (where $\phi = \pi$ is needed for negative angles) followed by a construction of the angular axis, conversion to intensity using Eq. (13), and then plotting.

line 87–98 Evaluation and plotting of intensities on the whole upper hemisphere using the same principle as above.

line 99–104 Function for calculating N Gauss points and weights on the interval $[a, b]$.

line 105–111 Function for calculating G using Eq. (19), where an approximation is used for large exponents to avoid overflow.

line 112–113 This ensures that `main()` is executed when the code is run.

C. Hankel Transform

The Hankel transform class is a minimalistic implementation of the MATLAB code from [22], which is based on the work found in [23]. For completeness, both a forward and inverse Hankel transform method are given even though the latter is never used. The math behind the code is explained in [22,23]; therefore only the usage is explained here.

line 1–9 Credits to the original paper, load of modules and special functions: `jn` is the Bessel function of order n, and `jn_zeros` returns the zeros of a Bessel function of order n.

line 10–13 Beginning of the `HankelTransform` class constructor, which takes the following arguments:

`Rmax`	max domain value for untransformed axis;
`order`	order of the Hankel transform;
`N`	number of (equidistant) discretization points.

line 14–16 Calculation of zeros for a Bessel function of order `order`. Calculation of the discretization of the radius axis in untransformed domain and similar for the transformed ("frequency") domain. The domains are stored so that they can be fetched by the user afterward; see line 49 and onward in the main file for an example of use.

line 17–24 Calculation of relevant matrices for the conversion; see [22,23] for a derivation of the results.

line 25–27 A function that returns the Hankel transform of order `order` of the function `f` given as an array calculated in the radius array points from line 14.

line 28–29 A function that returns the inverse Hankel transform of order `order` of the function `F` given as an array calculated in the frequency array points from line 15.

D. Modification and Comments

Since the implementation is kept simple, a few comments on how to modify the code for speed or usage are given here.

First of all, the Fresnel reflection from Eq. (14) is left out for simplicity. This is easily implemented by first calculating the Fresnel reflection for a given input angle and n_1, n_2 and afterward multiplying this with K just after line 70. Second, for repeated use, there is no reason to reinitialize the `HankelTransform` class more than once. Also, the Bessel function zeros calculated in the `HankelTransform` class can be precalculated and stored in a `numpy` (or MATLAB) data file, which is also shown in [22]. Other sizes that only should be calculated once for repeated use include `phi`, `wphi`, `mu`, `wmu`, and `rho`. All in all, the code is executed fast, and without visualization (but with initialization of all previously mentioned variables) a runtime of less than 0.5 s should be expected on a normal desktop computer.

For ACVs with a larger dynamic range than a Gaussian ACV, it is proposed [6] to use FFTlog or similar to obtain good precision.

The author finds that for a general implementation of the GHS method, a ρ, ϕ coordinate system is more appropriate than the current Cartesian discretization in the α, β plane [6,24]. See, for example, [24, Fig. 2], where only ~20% of the computational domain contains information, since the rest corresponds to "intensities" of complex angles ($\sqrt{\alpha^2 + \beta^2} > 1$). Furthermore it makes circular integrals more precise. For these reasons, a decrease in CPU time and memory consumption can be expected.

E. Verification

To verify the code, results from [6] for a rough Gaussian surface with Gaussian ACV were used. The number of Fourier series terms was put to 20, and μ was discretized in 100 points. The result is presented in Fig. 3(a) and is shown to match. This confirms that the code is implemented correctly. The result for 40° scattering angle in the complete ρ, ϕ plane is presented in Fig. 3(f).

It can in general be expected that the narrower the peak of the scattered energy, the more Fourier terms will be needed to accurately resolve the shape of the response, since higher frequencies are required. This is illustrated in Fig. 3(b), where the same result with $\theta_i = 40°$ is calculated for $N_f = 2$, 5, 10, 15, 99. Some of their corresponding ρ, ϕ plots are seen in Figs. 3(c)–3(f). From this it can be seen that using $N_f = 2$ gives a negative (hence infeasible) solution in some areas since the best approximation using only one cosine cannot counteract that effect. Using $N_f = 5$ still gives negative contributions even though that is not seen in the specular plot in Fig. 3(b). Already at $N_f = 10$, the result resembles the reference ($N_f = 99$) quite well, and for $N_f = 15$, the result is not distinguishable from the reference solution.

8. CONCLUSION

This paper presents a method for calculating rough surface reflectance based on the GHS method discretized using the DOM. The implementation is seen to reproduce the results from [6]. Since DOM requires a polar coordinate system, such is proposed and implemented in a way that exploits the isotropy assumption used for the rough surfaces under consideration. This makes it possible to efficiently use the GHS rough surface scattering method for radiative transfer problems. Furthermore it is shown how it is possible to keep the desired shift-invariance property from the direction cosine space in this new setting. It is also shown how the proposed discretization makes use of all data points in the discretized space as opposed to the approaches in [24,6].

A Python implementation is presented in Appendix A with the purpose of passing on a ready-to-use code in a freely available programming language, which hopefully will make the GHS method as accessible as possible, so that researchers can benefit from its improved accuracy compared to many other single scattering theories.

APPENDIX A: CODE

Here follows the code for a Python program to implement the above GHS method for polar coordinates. The source files are also available by e-mailing the author.

1. cosineHarveyShack.py

The following is the main file for the program. Note that it is dependent on `HankelLibSimple.py`.

```
#(c) Villads Egede Johansen, 2014, vejo@mek.dtu.dk
from __future__ import division
import numpy as np
import matplotlib.pyplot as plt

from numpy import cos,sin,arcsin,exp,pi,sqrt,multiply
try:    from numpy.polynomial.legendre import leggauss
except: from scipy.special.orthogonal import p_roots as leggauss
from HankelLibSimple import HankelTransform

def main():
   ############################
   # INPUT PARAMETERS
   lam = 10.6e-6
   sigmas = 2.27e-6
   lc = 20.9e-6
   thetai = -40 /180*pi
   n1 = 1.0
   n2 = -n1

   N = 100
   Nf = 20

   ############################
   # SETUP
   phi,wphi = calcGauss(2*N,0,pi)
   phi = phi[np.newaxis].T
   mu,wmu = calcGauss(N,0,1)
   rho = sqrt(1-mu**2)[::-1]

   f, wf = calcGauss(99,0,1/lam)
   fo = -sin(thetai)/lam
   fm = sqrt(f**2+fo**2-2*f*fo*(cos(phi)))
   PSD = pi* lc**2 *sigmas**2 *exp(-(pi*lc*fm)**2)
   sigmarel = sqrt(( (PSD*f).dot(wf).dot(2*wphi)))

   sigmashat = sigmas/lam
   sigmarelhat = sigmarel/lam
   lchat = lc/lam
   gammai = cos(thetai)

   rhom = sqrt(rho**2+sin(thetai)**2
         +2*rho*sin(thetai)*(cos(phi)))

   cosk =cos(multiply(np.arange(Nf),phi))
   sk = np.zeros((Nf,N))

   #Setup Hankel transform class
   ht = HankelTransform(Rmax=50,order=0,N=201)
   vrad,rrad = ht.v, ht.r

   ############################
   # Calculations
   for n in range(N):
      gammas = mu[::-1][n]
      A = exp(-(2*pi*(n1*gammai-n2*gammas)*sigmarelhat)**2)
      B = 1-A
      Cs = sigmashat**2 *exp(-rrad**2/lchat**2)
      G = calcG(gammai,gammas,sigmarel,sigmarelhat,sigmas,Cs,n1,n2)
      Srad = ht.fht(G)

      _rhom = rhom[:,n]
      idxNew = np.argsort(_rhom)
      rhoOrdered = _rhom[idxNew]
      idxBack = np.argsort(idxNew)
      Stot = B*np.interp(rhoOrdered,vrad,Srad)[idxBack]
      sk[:,n] = 2/pi*multiply(Stot,cosk.T).dot(wphi)

   K = 1/(pi*sk[0][::-1] *mu).dot(wmu)
   B = 1-exp(-(2*pi*(n1*gammai-n2*gammai)*sigmarelhat)**2)
   K = B*K

   ############################
   # VISUALIZATION
   Slinepos = 0.5*sk[0] + sum([ sk[k]*cos(k*0)
                                for k in range(1,sk.shape[0])])
   Slineneg = 0.5*sk[0] + sum([ sk[k]*cos(k*pi)
                                for k in range(1,sk.shape[0])])
   angles = arcsin(rho)
   angles = np.concatenate([-angles[::-1],angles])
   I = cos(angles) *K*np.concatenate([Slineneg[::-1],Slinepos])

   plt.plot(angles/pi*180,I,lw=2)
   plt.xlim(-90,90), plt.ylim(0)
   plt.show()

   phifull = np.linspace(0,2*pi,2*N+1)[np.newaxis].T
   cosk =cos(multiply(np.arange(Nf),phifull))
   f = 0.5*sk[0] + sum([ multiply(sk[k,:],cosk[:,k][np.newaxis].T)
                        for k in range(1,sk.shape[0])])
   f = mu[::-1]*K*f

   ax = plt.subplot(1,1,1, projection="polar", aspect=1.)
   cax=ax.pcolormesh(phifull.ravel(), rho, f.T)
   plt.gcf().colorbar(cax)
   plt.show()


def calcGauss(N,a,b):
   p,w = leggauss(N)
   p = (b-a)/2.*p + (b+a)/2.
   w = (b-a)/2.*w
   return p,w

def calcG(gammai,gammas,sigmarel,sigmarelhat,sigmas,Cs,n1,n2):
   exponent1 = (2*pi*(n1*gammai-n2*gammas)*sigmarel/sigmas)**2 *Cs
   exponent2 = (2*pi*(n1*gammai-n2*gammas)*sigmarelhat)**2
   if exponent2 > 50:
      return exp(exponent1-exponent2)
   return (exp(exponent1)-1)/(exp(exponent2)-1)

if __name__ == "__main__":
   main()
```

2. HankelLibSimple.py

The following is a minimalistic implementation of a class used to calculate Hankel transforms. The code is based on the approach presented in [23].

```
#(c) Villads Egede Johansen, 2014, vejo@mek.dtu.dk
#Port of matlab code by Manuel Guizar Sicairos based on
#  M. Guizar-Sicairos and J. C. Gutierrez-Vega, Computation of quasi-discrete
#  Hankel transforms of integer order for propagating optical wave fields,
#  J. Opt. Soc. Am. A 21, 53-58 (2004).
from __future__ import division
import numpy as np
from scipy.special import jn,jn_zeros

class HankelTransform:
   def __init__(self,Rmax,order=0,N=500):
      c = jn_zeros(order,N+1)

      self.r = c[:N] *Rmax/c[N]        #Radius vector
      self.v = c[:N] /(2*np.pi*Rmax)   #Frequency vector

      Jn,Jm = np.meshgrid(c[:N],c[:N])

      self._C = 2/c[N] *jn(order,Jn*Jm/c[N]) / (
            abs(jn(order+1,Jn)) * abs(jn(order+1,Jm)))

      self._m1 = abs(jn(order+1,c[:N]))/Rmax
      self._m2 = self._m1*Rmax/c[N] /(2*np.pi*Rmax)

   def fht(self,f):
      return np.dot(self._C,f/self._m1)*self._m2

   def iht(self,F):
      return np.dot(self._C,F/self._m2)*self._m1
```

ACKNOWLEDGMENTS

The author is grateful for the support from the Danish National Technology Foundation through the ODAAS project.

REFERENCES

1. I. Simonsen, "Optics of surface disordered systems," Eur. J. Phys. Special Topics **181**, 1–103 (2010).
2. J. R. Colin and A. A. Maradudin, *Light Scattering and Nanoscale Surface Roughness* (Springer, 2007).
3. J. E. Harvey, N. Choi, and A. Krywonos, "Scattering from moderately rough interfaces between two arbitrary media," Proc. SPIE **7794**, 77940V (2010).
4. J. E. Harvey, C. L. Vernold, A. Krywonos, and P. L. Thompson, "Diffracted radiance: a fundamental quantity in nonparaxial scalar diffraction theory," Appl. Opt. **38**, 6469–6481 (1999).
5. S. Schröder, A. Duparré, L. Coriand, A. Tünnermann, D. H. Penalver, and J. E. Harvey, "Modeling of light scattering in different regimes of surface roughness," Opt. Express **19**, 9820–9835 (2011).
6. A. Krywonos, J. E. Harvey, and N. Choi, "Linear systems formulation of scattering theory for rough surfaces with arbitrary incident and scattering angles," J. Opt. Soc. Am. A **28**, 1121–1138 (2011).
7. N. Choi and J. E. Harvey, "Numerical validation of the generalized Harvey–Shack surface scatter theory," Opt. Eng. **52**, 115103 (2013).
8. S. Faÿ, S. Dubail, U. Kroll, J. Meier, Y. Ziegler, and A. Shan, "Light trapping enhancement for thin-film silicon solar cells by roughness improvement of the ZnO front TCO," in *16th EC Photovoltaic Solar Energy Conference* (2000), pp. 361–364.
9. J. Stam, "An illumination model for a skin layer bounded by rough surfaces," in *Rendering Techniques 2001* (Springer, 2001), pp. 39–52.
10. A. B. Murphy, "Modified Kubelka–Munk model for calculation of the reflectance of coatings with optically-rough surfaces," J. Phys. D **39**, 3571–3581 (2006).
11. S. Schröder, A. Duparré, K. Füchsel, N. Kaiser, A. Tünnermann, and J. E. Harvey, "Scattering of roughened TCO films—modeling and measurement," in *Optical Interference Coatings* (2010).
12. M. Elias, P. Castiglione, and G. Elias, "Influence of interface roughness on surface and bulk scattering," J. Opt. Soc. Am. A **27**, 1265–1273 (2010).
13. S. Chandrasekhar, *Radiative Transfer* (Dover, 1960).
14. K. Stamnes, S.-C. Tsay, W. Wiscombe, and K. Jayaweera, "Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media," Appl. Opt. **27**, 2502–2509 (1988).
15. Z. Jin and K. Stamnes, "Radiative transfer in nonuniformly refracting layered media: atmosphere-ocean system," Appl. Opt. **33**, 431–442 (1994).
16. M. Elias, "Physics, colour and art: a fruitful marriage," J. Int. Colour Assoc. **8**, 25–35 (2012).
17. G. E. Thomas and K. Stamnes, *Radiative Transfer in the Atmosphere and Ocean* (Cambridge University, 2002).
18. Z. Jin, T. P. Charlock, K. Rutledge, K. Stamnes, and Y. Wang, "Modeling of light scattering from micro- and nanotextured surfaces," Appl. Opt. **45**, 7443–7455 (2006).
19. S. T. Thynell, "Discrete-ordinates method in radiative heat transfer," Int. J. Eng. Sci. **36**, 1651–1675 (1998).
20. W. J. Wiscombe, "The delta-M method: rapid yet accurate radiative flux calculations for strongly asymmetric phase functions," J. Atmos. Sci. **34**, 1408–1422 (1977).
21. O. Christensen, *Differentialligninger og uendelige ræ kker* (Institut for Matematik, Danmarks Tekniske Universitet, 2009).
22. M. Guizar, "Integer order Hankel transform," http://www.mathworks.com/matlabcentral/fileexchange/6570-integer-order-hankel-transform, 2004.
23. M. Guizar-Sicairos and J.-C. Gutiérrez-Vega, "Computation of quasi-discrete Hankel transforms of integer order for propagating optical wave fields," J. Opt. Soc. Am. A **21**, 53–58 (2004).
24. D. Dominé, F.-J. Haug, C. Battaglia, and C. Ballif, "Modeling of light scattering from micro- and nanotextured surfaces," J. Appl. Phys. **107**, 044504 (2010).

Publication P6

Reflectance Spectroscopy of Polyurethane-TiO2 Composite coating on Aluminium

Reflectance Spectroscopy of Polyurethane-TiO_2 Composite coating on Aluminium

Visweswara Chakravarthy Gudla[1], Villads Egede Johansen[1], Stela Canulescu[2], Jørgen Schou[21], Rajan Ambat[1*]

[1]Department of Mechanical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[2]Department of Photonics Engineering, Risø Campus, Technical University of Denmark, DK-4000 Roskilde, Denmark.

*Corresponding Author: ram@mek.dtu.dk, Phone: 0045-45252181

Contents

Abstract

The results of a physical simulation carried out using Polyurethane-TiO_2 composite coatings on bright and matte aluminium surfaces to understand the light scattering effect for designing decorative white anodised surfaces using aluminium matrix composites is presented. The idea is to incorporate light scattering particles in to the anodised layer for optical effects. Polyurethane (PU) matrix is selected for its matching refractive index (n=1.7) with anodic alumina layer. Three different TiO_2 particle size distributions were dispersed in PU and spin coated onto high gloss (HG) and caustic etched (CE) aluminium substrates. The reflectance spectra of PU-TiO_2 films were analysed using an integrating sphere-spectrophotometer. Data show that the PU-TiO_2 coatings have a high diffuse reflectance due to the multiple scattering from TiO_2 particles and coating substrate interface. The diffuse reflectance spectra of TiO_2 containing films varied weakly with particle concentration and reached a steady state value at 1 wt.%. Using Kubelka Munk two-stream model, the scattering and absorption coefficient of TiO_2 in PU was predicted. The studies presented in this paper provide valuable information on generating bright white decorative anodised aluminium surfaces starting with aluminium-TiO_2 composites.

Keywords: Anodised Aluminium, White, TiO_2, Reflectance, Caustic Etching, Gloss, Electrochemical Finishing

1. Introduction

Anodising is a widely used surface finishing technique applied to aluminium (Al) alloy components to improve their corrosion resistance, adhesion to paint top coats and pleasing aesthetics in architectural, decorative and automobile applications [1] [2] [3] [4]. The anodised Al layer is usually transparent to visible light and contains nano sized pores which can be filled with a variety of dyes to impart a wide spectrum of colours to the surface [5] [6] [7] [8] [9]. Nearly all shades of colours including black are imparted to anodised Al, but achieving a bright white and glossy anodised Al using conventional dyeing techniques has not been possible to date. The reason for this being the dye molecules used for colouring anodised Al are smaller than the pore sizes of the anodised layer (20-30 nm), whereas traditional white pigments are one order of magnitude larger [10]. Some anodising processes based on plasma electrolytic oxidation (PEO) or micro-arc oxidation (MAO) have been reported to generate white appearance due to hard ceramic aluminium oxide, but the surfaces are highly porous and diffuse with no gloss [11] [12] [13] [14] [15] [16] [17] [18] [19] [20]. Caustic etching on the other hand improves the scattering of light due to the surface roughness and also retains the surface gloss after anodising, but the reflectance is not as high as that of surfaces obtained from PEO/MAO and also there is a bluish metallic hue from the Al substrate [2] [3] [21] [22].

One alternative approach for achieving a white anodised aluminium surface is to anodise aluminium based composite that consists of oxide particles with required optical properties which are in contrast to those of anodic aluminium oxide. Also, the inertness of the oxide particles allows them to be embedded in the anodised layer during anodising process and cause diffuse scattering of light due to high refractive index together with a glossy appearance from a smooth anodised

surface. Using this approach, theoretically one could prepare white appearing anodised layer with high refractive index light scattering particles such as TiO_2 (n=2.6-2.9) [23] [24] [25] [26] similar to that in white paints. In order to do this, a fundamental understanding of how the thickness of the anodised layer, size and distribution of the TiO_2 particles affects scattering of light so that bluish metallic appearance from the substrate gets masked is required.

In this paper we study multiple scattering of light from TiO_2 particles embedded in a transparent medium, i.e. PU which mimics anodised Al in terms of refractive index (n=1.7) [27] [28] [29]. The PU-TiO_2 composites were coated onto Al substrates with different surface conditions. Scattering and absorption parameters are calculated using Kubelka-Munk two stream model for different particle size distributions. The effect of particle size distribution and amount of particles are experimentally investigated along with that of the initial Al surface condition. Reflectance measurements were performed using an integrating sphere-spectrophotometer setup that measures the total and diffuse reflectance of a coating. The TiO_2 powders were characterized for size, shape, and morphology using a laser particle analyser and a scanning electron microscope (SEM). Surface topography of Al substrates in terms of surface roughness was analysed with atomic force microscopy (AFM) [30]. Our studies serve as a model for understanding the interaction of light with high refractive index scattering particles in anodised aluminium matrix and facilitate designing of decorative anodised surfaces with required brightness and gloss.

2. Experimental

2.1. Materials & Methods

Aluminium (Al) substrates having high specular reflectance (high gloss) were obtained in cold rolled condition (Alcan Rolled Products, Germany). Substrates were subjected to ultrasonic etching in 10 wt. % NaOH at 60 ˚C for 10 min followed by demineralised water rinsing. The etched substrates were then desmutted in HNO_3 followed by demineralised water rinsing. A polyurethane (PU) clear coat based on acrylic resin (Sigmavar WS Satin™, Sigma coatings, the Netherlands) was used to mimic the anodic Al layer. TiO_2 particles in rutile phase of three different size distributions (DuPont Titanium Technologies, Belgium) were used. PU-TiO_2 dispersions with varying amount of TiO_2 (0.5, 1, 1.5, and 2 wt. %) were prepared by ultrasonic dispersion for 15 min. Coatings were deposited on as received high gloss and also on caustic etched Al substrates using a spin coater (Model WS-650Sz, Laurell Technologies Corporation, USA). Coating thickness of ~10 µm (measured using surface profilometry) was employed as most of the decorative anodising use an anodised layer thickness of ~10-15 µm [1]. The spin coated samples were dried in air. Nomenclature used for the prepared samples is given in Table 1.

Table 1: Nomenclature given for naming the prepared Al samples.

Substrate		**TiO_2 powder**		**Reflectance**	
Type	**Designation**	**Type**	**Designation**	**Type**	**Designation**
Aluminium - **H**igh **G**loss	**HG**	D_{50} – 320 nm	**320**	Total	**T**
Aluminium -**C**austic **E**tched	**CE**	D_{50} – 390 nm	**390**	Diffuse	**D**
		D_{50} – 500 nm	**500**		

A sample termed as ‘HG-320-0.5’ would mean that the coating contained 0.5 wt.% TiO_2 of size distribution D_{50} of 320 nm in PU over an as received high gloss Al substrate. Samples coated

with PU containing no TiO_2 were named as HG-PU-Ref and CE-PU-Ref. Samples without any coating were named as HG-Ref and CE-Ref.

2.2. Spectrophotometry

Optical appearance of the sample was analysed using an integrating sphere-spectrophotometer setup. The samples were illuminated with light from a Deuterium-Tungsten halogen light source (DH 2000, Ocean optics) at an angle of incidence of 8°. Reflected light was collected using a spectrometer (QE 65000, Ocean Optics). The integrating sphere allows total or diffuse reflectance measurements. The wavelength range analysed was 350-750 nm and the integration time was 4 s. The spectrometer was calibrated using NIST high diffuse alumina standard.

2.3. Characterization

The surface topography and roughness of the Al substrates in high gloss and caustic etched condition was measured using an atomic force microscope (Bruker Multimode 5). A silicon tip (Nanosensors SSS-NCH) with a tip radius of < 2 nm was used in tapping mode. The scan range was 100 x 100 μm^2. Morphology of the TiO_2 powder was observed using scanning electron microscopy (Model Quanta 200 ESEM FEG, FEI). Particle size analysis was performed on the TiO_2 powders using laser diffraction particle size analyser (Malvern Mastersizer 3000) coupled with a wet dispersion unit (Malvern Hydro SM).

2.4. Particle Size Distribution

Since agglomerates were present when measuring particle size distributions, the lower part of the particle size distributions were fitted to a lognormal distribution given by

$$N_p(r) = \frac{1}{r\beta\sqrt{2\pi}} e^{-(\ln(x)-\alpha)^2/(2\beta^2)},$$

where the two parameters α, β can be found from the mean particle size $\bar{r}$ and its variance σ^2 as $\beta^2 = \ln(1 + \sigma^2/\bar{r}^2)$ and $\alpha = \ln(\bar{r}) - \frac{1}{2}\beta^2$. The obtained results and the original data are shown in Figure 1. An overview plot comparing the different distributions are also seen. The 320 nm powder has a smaller standard deviation and a slightly shifted peak compared to the two others.

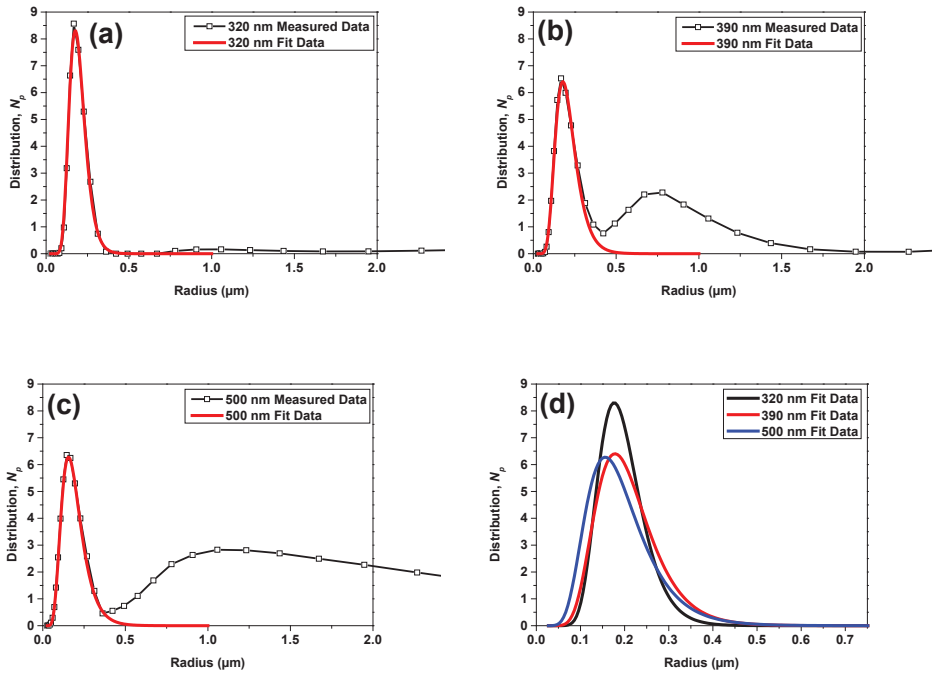


Figure 1: Lognormal fits and measured distributions for TiO_2 powder (a) 320 nm, $\bar{r} = 0.196, \sigma = 1.643$, (b) 390nm, $\bar{r} = 0.210, \sigma = 1.342$, (c) 500 nm, $\bar{r} = 0.194, \sigma = 1.117$, and (d) comparison of the different distributions. Especially for powders of 390 nm and 500 nm, the agglomerates are clearly visible.

3. Theory and Simulation

3.1. The Modified Kubelka-Munk model

In order to determine the best available particle distribution and best obtainable reflection under optimal conditions, a two-stream Kubelka-Munk model has been used. This model predicts diffuse and specular reflectivity for a model like the one in Figure 2. This model considers top and bottom surface reflections assuming plane interfaces and ascribes some effective scattering and absorption parameters to the PU-TiO_2 system which converts collimated light to diffuse light as it travels in the medium.

The calculations were carried out using the approach presented in [31] – which is a modification of the classical Kubelka-Munk (KB) model [32]. As shown in [31], this model fits better for measurements using an integrating sphere which will be used to characterize the final samples. One shortcoming of the model is that is does not take the roughness of the Al surface into account, but this is seen as a small loss compared to the simplicity of the model and the purpose of getting a general understanding of optimal parameters. From [31] the specular (R_s) and diffuse (R_d) reflections are calculated as

$$R_s = r_{cc}^{f},$$

$$R_d = \frac{(1-r_{cc}^{f})(1-r_{dd}^{b})R_{KM}}{1-r_{dd}^{b}R_{KM}},$$

where r_{cc}^{f} is the collimated-collimated reflectance from the air-PU interface, r_{cd}^{f} is the collimated-to-diffuse reflection from the air-PU interface, r_{dd}^{b} is the diffuse-diffuse reflection from the medium and back into the medium, and R_{KM} is the traditional KB surface reflection coefficient given by

$$R_{KM} = \frac{1-r_{dd}^{s}[a-b\coth(bSh)]}{a+b\coth(bSh)-r_{dd}^{s}}.$$

Here, S and K are respectively the effective scattering and absorption coefficients used in the KB model, $a = (S + K)/S$, $b = \sqrt{a^2 - 1}$, h is the height of the PU layer, and r_{dd}^{s} is the diffuse-diffuse reflection from the Al substrate back into the medium.

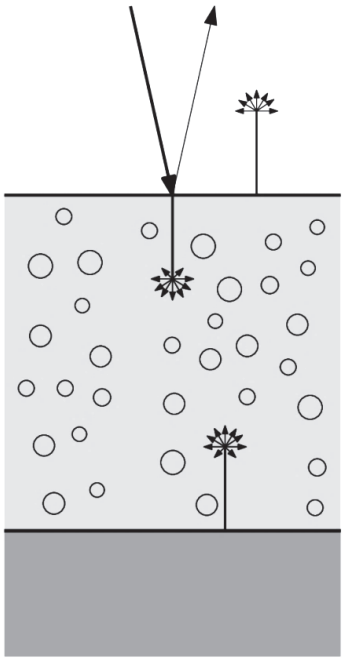


Figure 2: A model of an Al substrate with a PU-TiO_2 coating and plane interfaces. The incoming light (downwards thick arrow) is either reflected at the top surface or transported into the medium where it is scattered diffusely. The Al interface at the bottom reflects most of this diffuse light back, and in the end the diffuse light not absorbed leaves the PU coating again. The contribution from interface roughness is ignored in this model.

The reflection parameters are calculated using Fresnel reflection, and in particular the diffuse-diffuse reflection parameters are calculated as $r_{dd} = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} r(\theta)d\theta$, where r is the Fresnel reflection for the angle θ. This parameter is calculated using numerical integration and the results are seen in Figure 3(a).

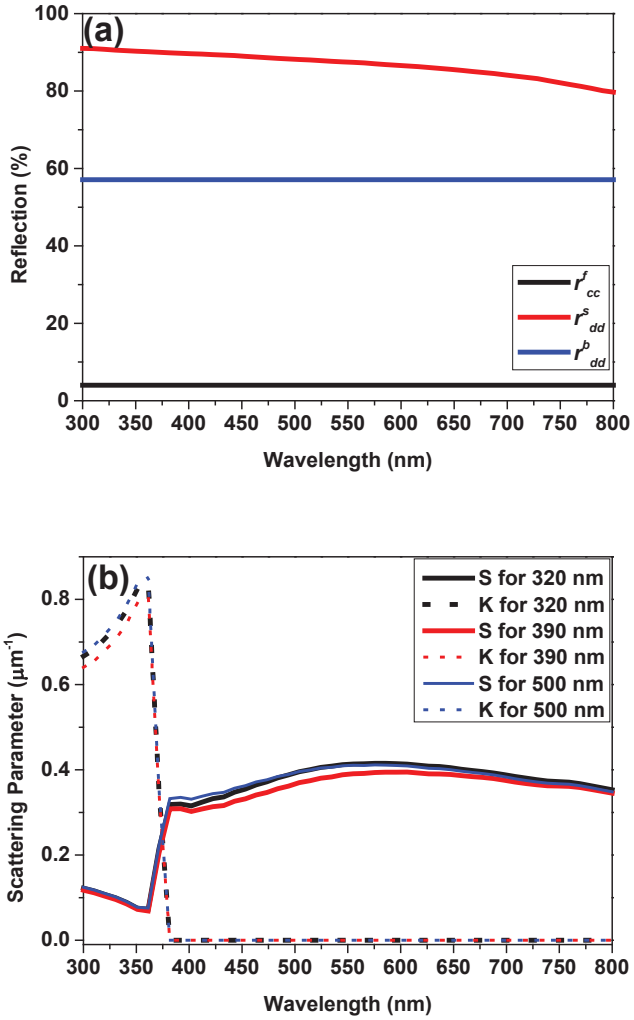


Figure 3: (a) Calculated reflection coefficients and (b) Calculated optical scattering parameters for different particle distributions (f = 10%).

3.2. Scattering and absorption parameters of particle distributions

Since S and K are heuristic parameters in the Kubelka-Munk model, there exist many ways of defining them, one method for spherical particles being [33]

$$S = \frac{3}{4} f \frac{C_{sca}(1-g)}{V_p},$$

$$K = 2f \frac{C_{abs}}{V_p},$$

where C_{sca}, C_{abs} are the particles scattering and absorption cross section respectively, f is the volume fraction of the particles in the PU, g is the so-called asymmetry parameter [34] and $V_p = \frac{4}{3}\pi r^3$ is the particle volume. Inherit in these expressions are (among others) that the particles act as single scatterers (they are optically far from each other), their size is in the Mie regime (which is controlled by the $\frac{3}{4}$ fraction), and that there is no depth (z) dependence on their scattering effect like seen in e.g. [35]. That the particles should be far from each other in turn means that the volume fraction, f, should be kept low as not to violate the validity of the scattering parameters calculation.

The parameters C_{sca}, C_{abs} can be calculated as explained in [34], and many computer implementations to do that exist. The freely available multiple spheres scattering tools MSTM 3.0 [36] was chosen. If only single particle scattering is of interest, then the programs BHMIE or BHCOAT that are based on code from [34] are available on the internet.

To be able to take size distribution of the particles into account, parameter calculation in [34] was generalized using the same procedure as in [33], which gives the average coefficients

$$< K >= 2f \frac{\int_0^\infty C_{abs}(r) N_p(r) dr}{\int_0^\infty V_p(r) N_p(r) dr},$$

$$< S >= f \frac{3}{4} \frac{\int_0^\infty C_{sca}(r)(1-g(r)) N_p(r) dr}{\int_0^\infty V_p(r) N_p(r) dr},$$

where r is the particle radius, and N_p is the normalized particle distribution, meaning that $\int_0^\infty N_p(r)dr = 1$. By using wavelength dependent optical data from [37] we calculated the S and K for the particle distributions and they can be seen in Figure 3 (b).

To confirm that particle material and size are well-suited for our experiments, a contour plot of S for different choices of refractive index and size in a resin with a refractive index (n) of 1.7 was generated. The result is presented in Figure 4, and clearly shows a peak in scattering efficiency matching our particle properties. It is furthermore seen, that by decreasing the particle size (or refractive index), the scattering efficiency rolls off fast, and it is therefore important to have a small variation in particle distribution.

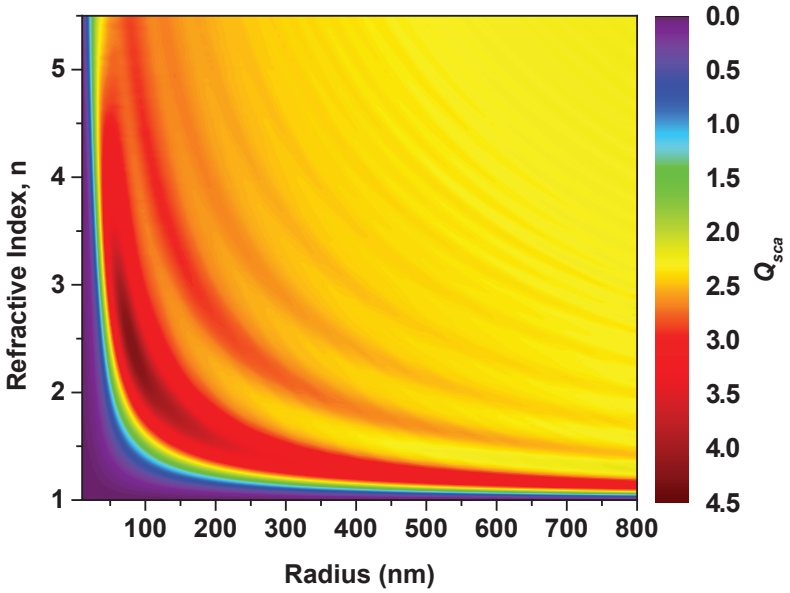


Figure 4: Scattering efficiency of a particle suspended in PU averaged over the visible spectrum.

3.3. Calculated Reflection

The above calculations show how r_{cc}^{f}, r_{ad}^{b}, r_{ad}^{s} are fixed due to the material properties of PU, and furthermore that *S*, *K* can be varied by choosing between our available particle distribution. Furthermore, the height of the coating, *h* and the particle volume fraction, *f*, is free parameters as well. We therefore studied total reflection for varying cases of these parameters.

Since, *K* for all particle distributions is close to zero in the visible light range (~380 nm-750 nm); it does not influence visual appearance much. However, the scattering coefficient does influence the visual appearance and the 320 nm distribution is seen to be distinctly better than the two other powder size distributions used. For the overall reflection it was found that the difference is much less pronounced, and the three distributions are close to being of same quality. Therefore we only present results for the 320 nm distribution. By varying the height *h* in the modified Kubelka-Munk model and using the data presented above, the reflections presented in Figure 5 were obtained. The results show first of all that an almost flat spectral reflection should be obtainable and furthermore that a layer much thicker than traditional oxide thicknesses (~10−15 μm) is required for obtaining a good white reflection. From the formulas it is seen that the effect of varying volume fraction and height by a factor is the same. This means that if *f* = 20% was aimed at only half the heights from the plot would be needed to obtain the same effect. From this it can be concluded that a high volume fraction - still satisfying the single scattering assumption - is needed in order to obtain a good scattering effect within a reasonable height.

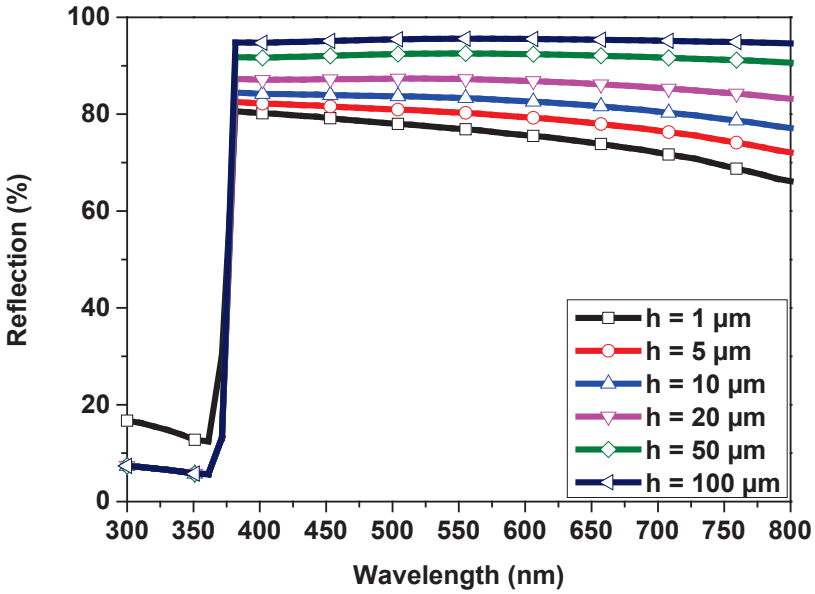


Figure 5: Calculated total reflectance (R_s+R_d) for different heights of an ideally coated sample with a filling fraction of f = 10% for TiO_2 powder of 320 nm size distribution.

4. Experimental Results

4.1. Scanning electron microscopy

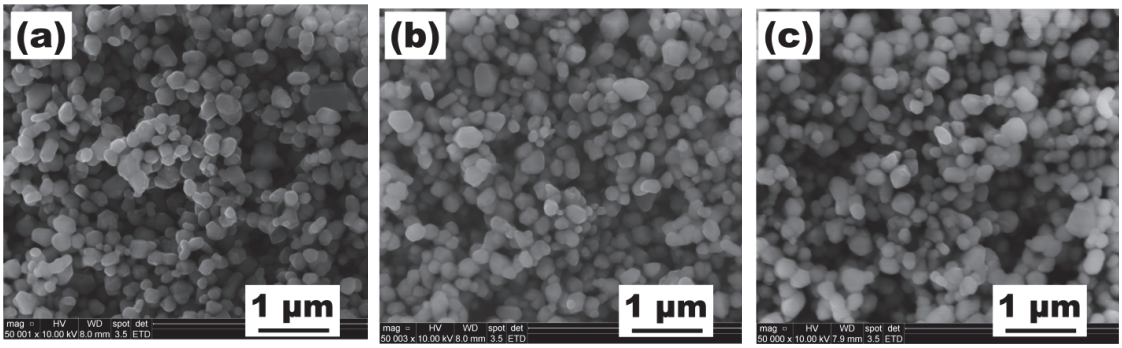


Figure 6: SEM micrographs of (a) TiO_2 powders-320 nm, (b) TiO_2 powders-390 nm and (c) TiO_2 powders-500 nm size distribution.

The SEM images of TiO_2 powders used for preparing the composite coatings are shown in **Figure 6**. It can be seen that the particles are spherical in shape and have a homogenous morphology. The size distribution (D_{50} value) measured using laser diffraction particle size analysis was found to be 320, 390, and 500 nm (Figure 1).

4.2. Surface morphology

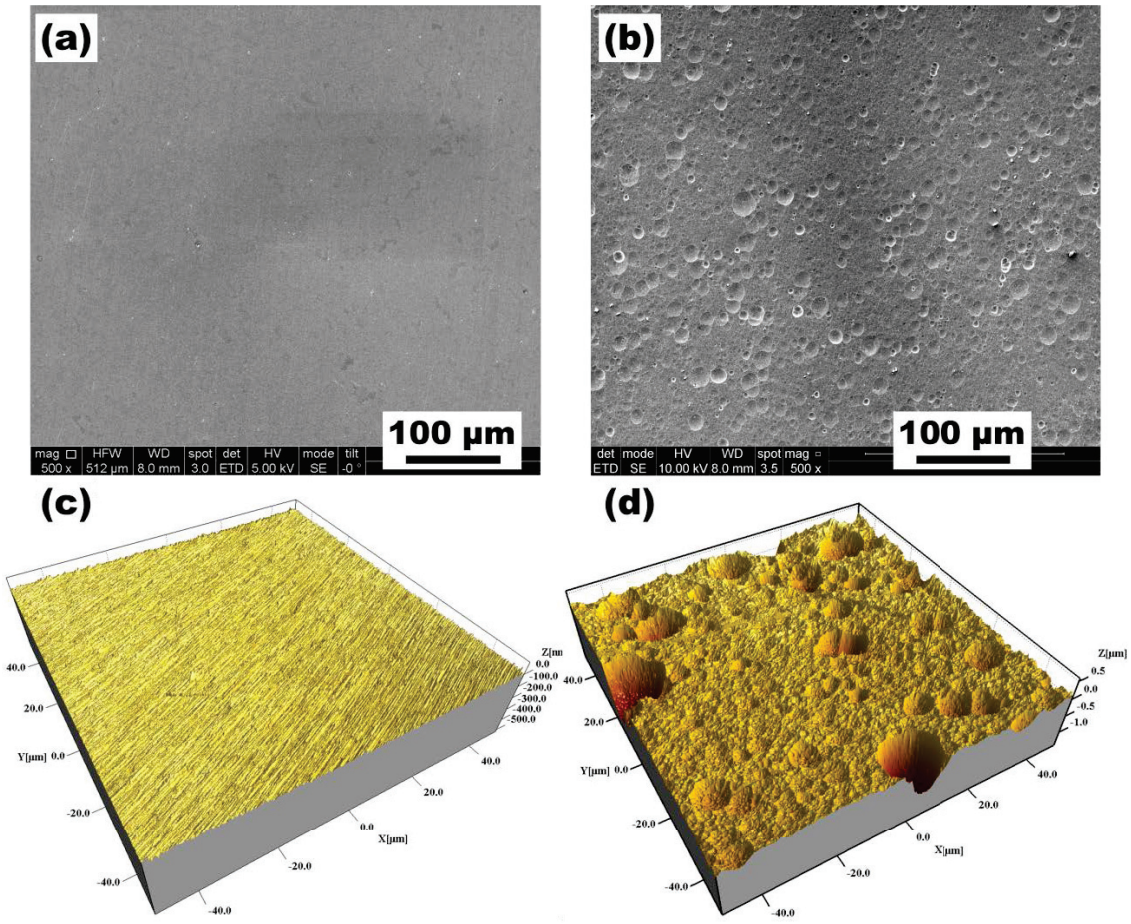
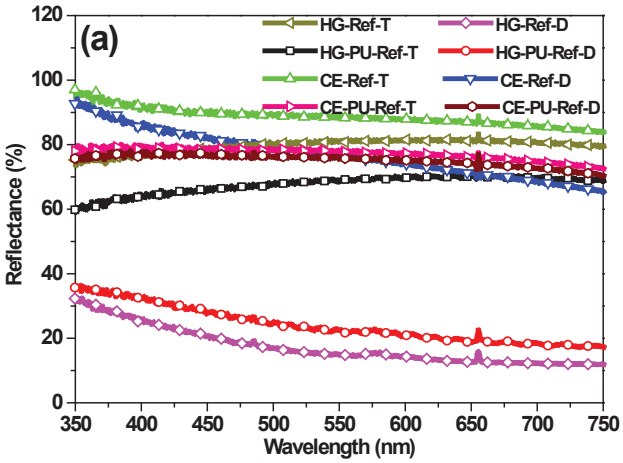


Figure 7: SEM micrographs of (a) high gloss aluminium, (b) caustic etched aluminium and AFM surface topography images of (c) high gloss aluminium and (d) caustic etched aluminium.

The surface of as received high gloss Al substrate as shown in the SEM and AFM images (**Figure 7** (a) & (c)) shows a typical rolled surface appearance with surface streaks in the rolling direction. The caustic etched aluminium surface (**Figure 7** (b) & (d)) shows a homogenous attack by the etching agent which has resulted in the formation of hemi-spherical pits/scallops on the surface. Also, it can be seen that the prior rolled morphology of the surface is lost after caustic etching. Area roughness (S_a, measured from AFM topography data) of the high gloss substrate was measured to be ~8 nm which after caustic etching treatment increased to a value of ~158 nm.

4.3. Spectrophotometry

4.3.1. Coatings without TiO_2



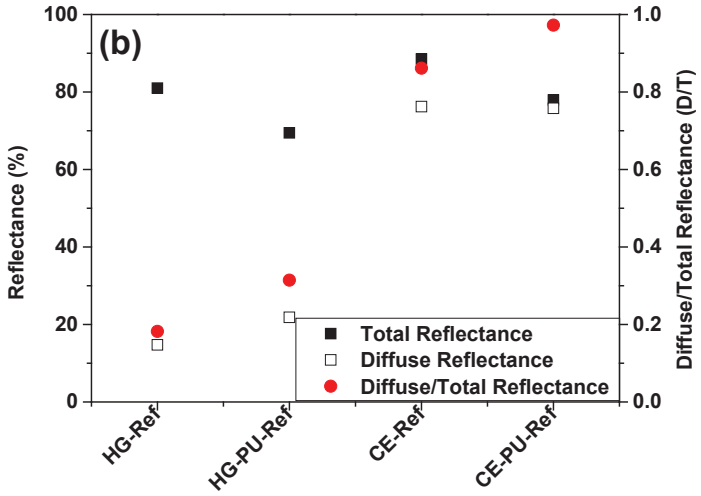


Figure 8: (a) Reflectance spectra and (b) respective intensity values at the wavelength of 555 nm of aluminium substrates (high gloss, HG and caustic etched, CE) with and without PU coating and D/T ratio (D- Diffuse and T-Total reflectance).

The reflectance spectra of the two different Al substrates with and without PU coating showed little intensity dependence on the wavelength (**Figure 8** (a)), except for slight intensity dependence at lower wavelengths. For the high gloss substrates, the reflectance was lower than average at lower wavelengths, but for etched substrate it was higher than average at lower wavelengths. Hence, the reflectivity of each coating at wavelength of 555 nm, where the human eye is most sensitive is summarized in **Figure 8** (b). High gloss aluminium substrate (HG-Ref) has a total reflectance (**Figure 8** (b)) of about 80% which upon etching increases to 90% (CE-Ref). The diffuse reflectance of high gloss aluminium is about 15% only whereas that for the caustic etched sample is 75%. The total and diffuse reflectance of high gloss substrates after being coated with PU (HG-PU-Ref) is 70% and 20%, respectively indicating that the PU coating reduces the reflectance of the sample and increases diffuse scattering of light. The reflectivity of the PU coating on caustic etched substrate reduces the total reflectance value to 80%, but does not affect the diffuse

reflectance considerably. As can be seen from the diffuse to total reflectance ratio (right axis in **Figure 8** (b)), the most diffuse appearing substrate surfaces are caustic etched with PU coating, but the brightest is caustic etched surface. The glossiest appearing surface is the as received HG aluminium surface.

4.3.2. Coatings with TiO_2

The reflectance spectra obtained from surfaces after coating with PU-TiO_2 showed absorption around 350-380 nm due to the TiO_2. There is no intensity dependence on wavelength above this value up to 750 nm. Representative total reflectance spectra obtained from PU-TiO_2 coatings of size distribution 390 nm on high gloss substrates is shown in Figure 9 for reference.

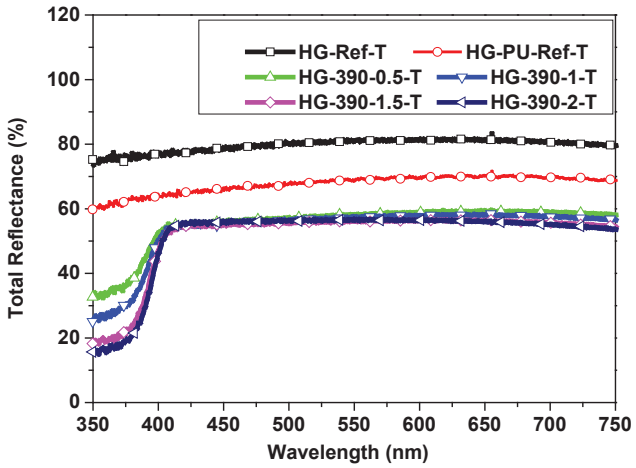


Figure 9: Total reflectance spectra of PU-TiO_2 (390 nm) coatings on high gloss substrates with varying TiO_2 content.

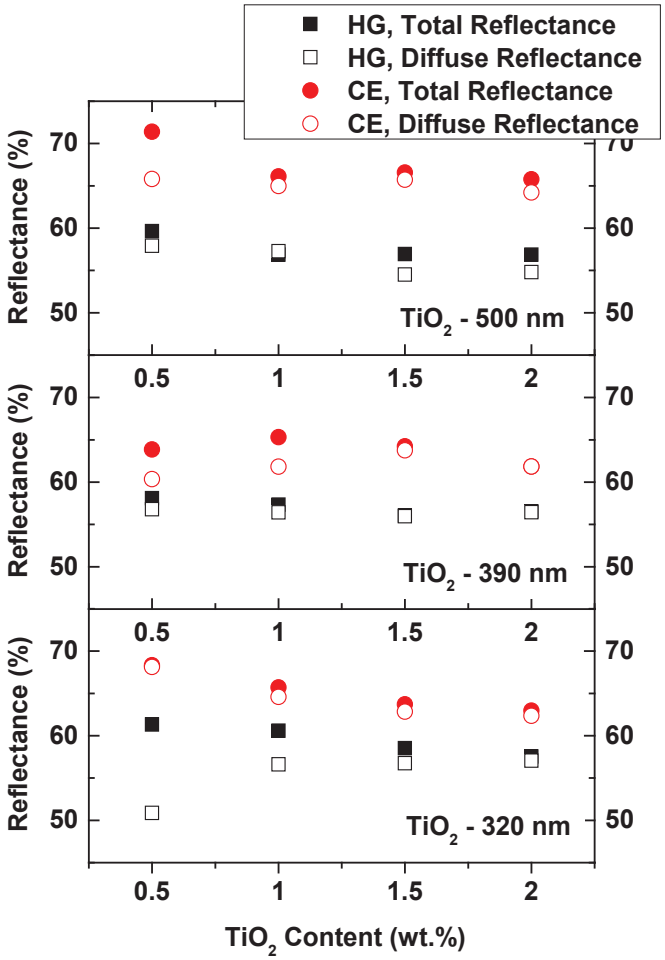


Figure 10: Reflectance values at 555 nm obtained for high gloss (HG) and caustic etched (CE) Al substrates coated with PU containing TiO_2 powders of different size distributions.

The total and diffuse reflectance values at the wavelength of 555 nm for PU-TiO_2 coatings of different TiO_2 size distributions on high gloss and caustic etched Al substrates is shown in **Figure 10**. For high gloss substrates (HG), the total reflectance shows a very slight decrease with increasing TiO_2 content for all the powder size distributions. The diffuse reflectance however, shows a decreasing trend with increasing powder content for powders of 390 nm and 500 nm, while increases for powder of 320 nm. For caustic etched surfaces (CE), the total reflectance value

decreases with increasing TiO_2 for powders 320 nm and 500 nm, while it shows a maximum at 1 wt.% for powder 390 nm. The diffuse reflectance decreases with increasing TiO_2 for powder 320 nm and is not intensely affected for powders 390 nm and 500 nm.

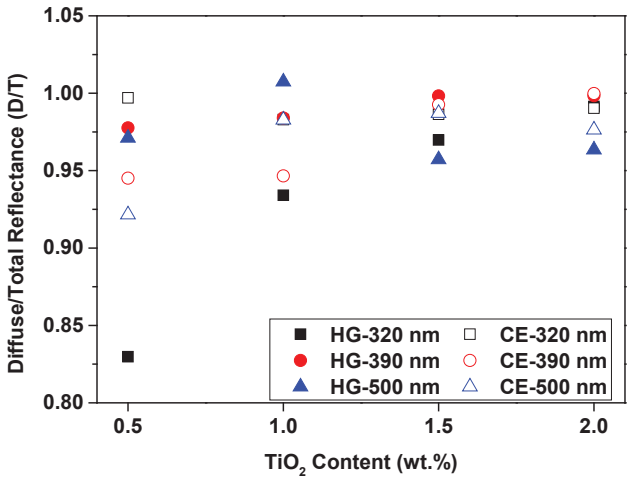


Figure 11: Diffuse to total reflectance ratio of the Al substrates coated with PU containing TiO_2 powders of different size distributions.

The ratio of diffuse to total reflectance (D/T) is presented in **Figure 11** for PU coatings with different powders on high gloss as well as caustic etched substrates. For coatings containing 320 nm sized powders on high gloss substrates, the D/T ratio increases with increasing TiO_2 content. For caustic etched substrates, the D/T ratio does not vary significantly with increasing TiO_2. For powders of 390 nm, the D/T ratio increases with increasing TiO_2 on high gloss as well as caustic etched substrates. Saturation is observed for caustic substrates at 1.5 wt.% TiO_2. For 500 nm powder, the D/T ratio does not show any observable trend on high gloss substrates, but increases with increasing TiO_2 content on caustic etched substrates. Of all the coatings prepared the most diffuse appearing on high gloss substrates is for 500 nm powder at 1 wt.% (HG-500-1) and on caustic etched substrates is for 320 nm powder at 0.5 wt.% (CE-320-0.5) and 390 nm at 2 wt.%

(CE-390-2). The least diffuse appearing surfaces are 320 nm powder at 0.5 wt.% on high gloss (HG-320-0.5) and 500 nm powder at 0.5 wt.% on caustic etched substrates (CE-500-0.5).

5. Discussion

The theoretical calculations show that the best scattering efficiency for a single sphere in PU is achieved when the particle size has a diameter of ~190-240 nm and a refractive index of ~2.2-2.9 (Figure 4). The efficiency of scattering and hence reflectance also increases with increasing thickness (Figure 5). However, an anodised layer of thickness higher than 15 μm is not feasible for decorative applications due to mechanical integrity issues [2] [38]. A lower film thickness (of ~10 μm) would require more scattering centres implying a larger particle concentration which might lead to optical crowding. For the powder size distributions used for preparing the coatings in the present study, theoretical calculations (Figure 3 (b)) show that powder of 320 nm gives the best scattering, while powders of 500 nm and 390 nm behave similarly and follow closely. This is in accordance with the experimentally measured reflectance values where powder of 320 nm gives highest scattering followed closely by powders of 390 nm and 500 nm (see **Figure 10**).

The measured reflectance values of PU- TiO_2 coatings on high gloss substrates are about 60 %. However, caustic etching treatment of the Al substrates increases the total reflectance and decreases the specular reflectance of the PU-TiO_2 coatings. The highest reflectance measured is 70 % and is for the coatings on caustic etched substrate. Also, the most diffuse appearing surfaces are achieved by caustic etching of the surfaces and at 2 wt.% of 390 nm powder (CE-390-2) and at 0.5 wt.% for 320 nm powder (CE-320-0.5).

Theoretical calculations based on Kubelka Munk two-stream model predict an enhancement of the diffuse reflectance with increasing thickness or particle concentration (see Figure 5). This trend is observed for particles which exhibit the largest scattering efficiency (320 nm) coated on highly reflective substrates. Indeed, the largest reflectance of the TiO_2-PU coating on high gloss substrates is observed at a concentration of 2 wt. %. However, for 390 nm and 500 nm TiO_2 powders, the diffuse reflectance decreases slightly with increasing particle concentration. A reason for this trend can be the agglomeration of the TiO_2 particles in the composite coatings. Agglomerated TiO_2 particles behave as a single larger particle rather than a smaller one, and the scattering efficiency is therefore moved away from the optimum [10]. Also, concentration of TiO_2 in the prepared composites that are higher than the critical pigment volume concentration would lead a reduced gloss and contribute to lower reflectance [10] [39] [40].

White anodising of Al using PEO and MAO techniques have typically reflectance values up to 80% [2], but the resulting anodised surfaces are totally diffuse. On the other hand, the physically simulated coatings prepared on the etched Al substrates in the current study have a reflectance of ~70 % with a specular component of about 5%. The PU coating when applied onto the high gloss and caustic etched substrates shows a reduction in the total reflectance value by 10-15 percentage points (**Figure 8**). This reduction in reflectance would however not be the case for decorative anodised pure Al as the anodic layer shows reduction in reflectance of only up to 5 percentage points [29] [41]. So, under the optimum size, distribution and concentration of TiO_2 particles, the reflectance for anodised Al with light scattering particles would be expected to be 5-10 percentage points higher (approx. 75-80 % total reflectance) than the measured values for the PU-TiO_2 coated surfaces. So, for anodic alumna with TiO_2 powders of 390 nm on a caustic etched substrate the total reflectance would be about 80% (with more than 90% of it as diffuse component). This reflectance

value is comparable to the surfaces obtained by PEO or MAO techniques and also is suitable for decorative applications due to the specular component which gives the surface gloss.

6. Conclusions

- Kubelka-Munk two stream model can serve as an effective tool to simulate and estimate the optical reflectance of anodic alumina containing light scattering centres.
- The TiO_2 powders of various size distributions dispersed in Polyurethane coated over Al substrates can be used to tailor the optical appearance of anodised aluminium due to the similarity in refractive index.
- Caustic etching of Al substrates increases the diffuse reflectance of the substrates, while PU coating reduces the total reflectance and increases the diffuse reflectance. Increase in the amount of light scattering TiO_2 particles decreases the specular reflectance thus making the surface appear more diffuse.
- The TiO_2 particle size distribution for achieving the maximum diffuse reflectance in the current tested range is observed to be 2 wt.% for PU-TiO_2 (390 nm) on caustic etched Al, but theoretical calculations predict an even higher content of TiO_2 is required.

Acknowledgements

The authors would like to thank Kai Dirscherl, Danish Fundamental Metrology, for help with the AFM measurements. Mette Larsen, Chemical Engineering Dept., DTU is thanked for help with the particle size analysis. The Danish National Advanced Technology Foundation is gratefully

acknowledged for the financial funding. The ODAAS project partners are acknowledged for their help.

References

[1] S. Mohagheghi, A. Hatefi, A. Kianvash, Effect of H3PO4 anodising variations on Al-epoxy adhesion strength, Surf. Eng. 29 (2013) 737–742. doi:10.1179/1743294413Y.0000000158.

[2] S. Wernick, R. Pinner, P.G. Sheasby, A.S.M. International., The surface treatment and finishing of aluminium and its alloys 1-2, Finishing Publ, Teddington, 1987.

[3] C. Grubbs, Decorative and architectural anodizing, Met. Finish. (1995).

[4] H.-J. Spies, Surface engineering of aluminium and titanium alloys: an overview, Surf. Eng. 26 (2010) 126–134. doi:10.1179/174329409X451146.

[5] C.A. Grubbs, Anodizing of aluminum, Met. Finish. 105 (2007) 397–412.

[6] Thompson, Wood, Porous anodic film formation on aluminium, Nature. 290 (1981) 230–232.

[7] P.G. Sheasby, The weathering of anodized aluminium, Alum. 2000. (1990).

[8] M. Selvam, Colouring of anodised aluminium by electroless method, Surf. Eng. 27 (2011) 711–718. doi:10.1179/1743294411Y.0000000022.

[9] H. Rezagholi, K. Zangeneh-Madar, M. Mirjani, M. Ahangarkani, Study on electropolished/anodised aluminium foil before and after dyeing, Surf. Eng. 30 (2014).

[10] DuPont ^{TM} Ti-Pure titanium dioxide, (n.d.).

[11] C. Siva Kumar, S.M. Mayanna, K.N. Mahendra, A.K. Sharma, R. Uma Rani, Studies on white anodizing on aluminum alloy for space applications, Appl. Surf. Sci. 151 (1999) 280–286. doi:http://dx.doi.org/10.1016/S0169-4332(99)00290-1.

[12] C. Siva Kumar, A.K. Sharma, K.N. Mahendra, S.M. Mayanna, Studies on anodic oxide coating with low absorptance and high emittance on aluminum alloy 2024, Sol. Energy Mater. Sol. Cells. 60 (2000) 51–57. doi:http://dx.doi.org/10.1016/S0927-0248(99)00062-8.

[13] W. Kaysser, Surface modifications in aerospace applications, Surf. Eng. 17 (2001) 305–312. doi:10.1179/026708401101517926.

[14] A. Rajendra, B.J. Parmar, A.K. Sharma, H. Bhojraj, M.M. Nayak, K. Rajanna, Hard anodisation of aluminium and its application to sensorics, Surf. Eng. 21 (2005) 193–197. doi:10.1179/174329405X50000.

[15] W. Xue, Z. Deng, R. Chen, T. Zhang, Microstructure and mechanical properties near interface between microarc oxidation coating and Al alloy substrate, Surf. Eng. 16 (2000) 344–348.

[16] J. George, N.J. Babu, Electron beam evaporated rigid optical solar reflectors for space applications, Surf. Eng. 21 (2005) 254–256. doi:10.1179/174329405X50127.

[17] A.K. Sharma, Surface engineering for thermal control of spacecraft, Surf. Eng. 21 (2005) 249–253. doi:10.1179/174329405X50118.

[18] Effect of current frequency on coating properties formed on aluminised steel by plasma electrolytic oxidation., Surf. Eng. 30 (2014).

[19] A.L. Yerokhin, X. Nie, A. Leyland, A. Matthews, S.J. Dowey, Plasma electrolysis for surface engineering, Surf. Coatings Technol. 122 (1999) 73–93. doi:10.1016/S0257-8972(99)00441-7.

[20] J.A. Curran, T.W. Clyne, Thermo-physical properties of plasma electrolytic oxide coatings on aluminium, Surf. Coatings Technol. 199 (2005) 168–176. doi:http://dx.doi.org/10.1016/j.surfcoat.2004.09.037.

[21] S. Bagherifard, M. Guagliano, Review of shot peening processes to obtain nanocrystalline surfaces in metal alloys, Surf. Eng. 25 (2009) 3–14. doi:10.1179/026708408X334087.

[22] R. Priestner, D.M. Priestner, Importance of substrate in surface engineering, Surf. Eng. 7 (1991) 53–59.

[23] U. Diebold, The surface science of titanium dioxide, Surf. Sci. Rep. 48 (2003) 53–229. doi:10.1016/S0167-5729(02)00100-0.

[24] J. Winkler, Titanium dioxide, Vincentz Verlag, Hannover, 2003.

[25] J.K. Yao, H.L. Huang, J.Y. Ma, Y.X. Jin, Y.A. Zhao, J.D. Shao, et al., High refractive index TiO2 film deposited by electron beam evaporation, Surf. Eng. 25 (2009) 257–260. doi:10.1179/026708408X329498.

[26] W.X. Dai, X. Chen, E. Li, X.X. Wang, P. Liu, X.Z. Fu, Influence of pH value of TiO2 sol on surface gloss of corresponding TiO2 film coated on ceramic tiles, Surf. Eng. 25 (2009) 106–110. doi:10.1179/174329408X326498.

[27] I.H. Khan, J.S.L. Leach, N.J.M. Wilkins, The thickness and optical properties of films of anodic aluminium oxide, Corros. Sci. 6 (1966) 483–497. doi:http://dx.doi.org/10.1016/S0010-938X(66)80055-0.

[28] G.F. Pastore, Transmission interference spectrometric determination of the thickness and refractive index of barrier films formed anodically on aluminum, Thin Solid Films. 123 (1985) 9–17. doi:http://dx.doi.org/10.1016/0040-6090(85)90036-7.

[29] S. Canulescu, K. Rechendorff, C.N. Borca, N.C. Jones, K. Bordo, J. Schou, et al., Band gap structure modification of amorphous anodic Al oxide film by Ti-alloying, Appl. Phys. Lett. 104 (2014) 121910. doi:10.1063/1.4866901.

[30] J.T. Thornton, Metal surface investigation by scanning probe microscopy, Surf. Eng. 16 (2000) 287–293.

[31] A.B. Murphy, Modified Kubelka–Munk model for calculation of the reflectance of coatings with optically-rough surfaces, J. Phys. D. Appl. Phys. 39 (2006) 3571.

[32] Kubelka, Munk, Reflection characteristics of paints, Zeitschrift Fur Tech. Phys. 12 (1931) 593–601.

[33] J.C. Auger, R.G. Barrera, B. Stout, Scattering efficiency of clusters composed by aggregated spheres, J. Quant. Spectrosc. Radiat. Transf. 79-80 (2003) 521–531.

[34] C.F. Bohren, D.R. Huffman, Absorption and Scattering of Light by Small Particles, Wiley-VCH, 1998.

[35] W.E. Vargas, G.A. Niklasson, Generalized method for evaluating scattering parameters used in radiative transfer models, J. Opt. Soc. Am. A. 14 (1997) 2243–2252.

[36] http://www.eng.auburn.edu/users/dmckwski/scatcodes/, (n.d.).

[37] E.D. Palik, Chapter 2 - Refractive Index, Handbook of Optical Constanst of Solids, (1997) 5–114. doi:http://dx.doi.org/10.1016/B978-012544415-6.50149-7.

[38] C. Grubbs, Anodizing of aluminum, Met. Finish. (1999) 397–412.

[39] W.K. Asbeck, M. Van Loo, Critical Pigment Volume Relationships., Ind. Eng. Chem. 41 (1949) 1470–1475. doi:10.1021/ie50475a042.

[40] G.P. Bierwagen, D.C. Rich, The critical pigment volume concentration in latex coatings, Prog. Org. Coatings. 11 (1983) 339–352. doi:http://dx.doi.org/10.1016/0033-0655(83)85007-9.

[41] S. Van Gils, P. Mast, E. Stijns, H. Terryn, Colour properties of barrier anodic oxide films on aluminium and titanium studied with total reflectance and spectroscopic ellipsometry, Surf. Coatings Technol. 185 (2004) 303–310. doi:http://dx.doi.org/10.1016/j.surfcoat.2004.01.021.

DTU Mechanical Engineering
Section of Solid Mechanics
Technical University of Denmark

Nils Koppels Allé, Bld. 404
DK- 2800 Kgs. Lyngby
Denmark
Phone (+45) 4525 4250
Fax (+45) 4593 1475
www.mek.dtu.dk
ISBN: 978-87-7475-395-7